

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 12:58:43 ; Search time 85 Seconds  
(without alignments)  
1641.714 Million cell updates/sec

Title: US-09-775-803-12\_COPY\_1\_389  
Perfect score: 1938  
Sequence: 1 MLRSALLSAVLALLRAQFPF.....LRHNRLRALPRTLFRNLSSL 389

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1938	100.0	566	3	AAY69183 Amino aci
2	1938	100.0	566	5	AAU78891 Mouse end
3	1477	76.2	560	2	AAR71294 Human gly
4	1477	76.2	560	3	AAY69184 Amino aci
5	1477	76.2	560	5	AAU78892 Human end
6	1477	76.2	560	7	ADF69095 Human MP5
7	607	31.3	581	5	ABB53264 Human pol
8	607	31.3	581	6	ABU52615 Human NOV
9	607	31.3	581	6	ABU56706 Lung canc
10	607	31.3	581	6	ABU56544 Lung canc
11	607	31.3	581	6	ABR43175 Human REM
12	607	31.3	581	7	ADC83709 Human leu
13	607	31.3	581	7	ADE08191 Novel pro
14	607	31.3	581	7	ADN39956 Cancer/an
15	607	31.3	581	7	ADN39092 Cancer/an
16	607	31.3	790	7	ADC86793 Human GPC
17	607	31.3	819	6	ABP97224 Tumour-as
18	583	30.1	578	7	ADC83708 Rat leuci
19	547	28.2	444	7	ADE07921 Novel pro
20	509.5	26.3	395	6	ABU12114 Human pro
21	503.5	26.0	570	8	ADI36915 Human LRR
22	445	23.0	545	5	ABP64705 Human pro
23	445	23.0	545	6	ABJ26653 Human pro
24	445	23.0	545	6	ABU54560 Human NOV
25	445	23.0	545	6	ABU12117 Human pro

26	442	22.8	545	6	ABU54561 Human NOV
27	419.5	21.6	605	2	AAR85888 WD-40 dom
28	419.5	21.6	605	6	AAO23109 IGFALS 'h
29	419.5	21.6	605	7	ADD48887 Human Pro
30	419.5	21.6	605	7	ADD47835 Human Pro
31	419.5	21.6	605	7	ADE61677 Human Pro
32	419.5	21.6	605	7	ADJ68826 Human hea
33	382.5	19.7	603	2	AAR85889 WD-40 dom
34	382.5	19.7	603	7	ADD48824 Rat Prote
35	382.5	19.7	603	7	ADD48885 Rat Prote
36	382.5	19.7	603	7	ADE61675 Rat Prote
37	382.5	19.7	603	7	ADD47833 Rat Prote
38	379.5	19.6	928	8	AAO24336 Human Sli
39	379.5	19.6	1122	8	AAO24335 Human Sli
40	379.5	19.6	1461	8	AAO24327 Human Sli
41	379.5	19.6	1508	2	AAY27141 Human sli
42	379.5	19.6	1508	2	AAW96706 Protein s
43	379.5	19.6	1508	2	AAY04138 Human sli
44	379.5	19.6	1531	7	ADG42633 Rat SLIT
45	379.5	19.6	1534	2	AAW46966 Amino aci

ALIGNMENTS

RESULT 1  
AAY69183  
ID AAY69183 standard; protein; 566 AA.  
XX  
AC AAY69183;  
XX

DT 30-MAY-2000 (first entry)  
XX

DE Amino acid sequence of murine glycoprotein V (GP V).  
XX

KW Glycoprotein V; GP V; transgenic animal; platelet; thrombin substrate;  
KW platelet function; haemostasis; platelet aggregation; anticoagulant;  
KW thrombolytic; antiplatelet therapeutic drug.  
XX

OS Mus sp.  
XX

FH Key Location/Qualifiers  
XX

FT Misc-difference 566  
FT /note= "encoded by TG"  
XX

PN WO200008137-A2.  
XX

PD 17-FEB-2000.  
XX

PF 04-AUG-1999; 99WO-US017594.  
XX

PR 04-AUG-1998; 98US-0109797P.  
XX

PA (CORT-) COR THERAPEUTICS INC.  
XX

PI Ramakrishnan V, Phillips DR;  
XX

DR WPI; 2000-195570/17.  
XX

DR N-PSDB; AAZ61216.  
XX

PT Transgenic animals with a modified glycoprotein V gene used to identify  
agents that modulate the effects of glycoprotein V.  
XX

PS Example 1; Fig 2; 44pp; English.  
XX

CC The present sequence represents a murine glycoprotein V (GP V). The  
specification describes non-human transgenic animals which contain or  
comprise a modified GP V gene. GP V is a platelet and endothelial cell  
specific glycoprotein, and is a substrate for thrombin. Activation of  
platelets by thrombin results in the loss of surface GP V. Platelets from  
the transgenic animals can be used in a number of assays to identify  
agents that modulate GP V function, or to assess the role of GP V in  
platelet function. Such assays may help elucidate the extent to which GP

CC V is critical for normal haemostasis. The transgenic animals may also be  
CC used to identify agents that modulate platelet aggregation. Agents that  
CC can be tested include anticoagulants, thrombolytics, and antiplatelet  
CC therapeutic drugs  
XX  
SQ Sequence 566 AA;

Query Match 100.0%; Score 1938; DB 3; Length 566;  
Best Local Similarity 100.0%; Pred. No. 2e-170;  
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSALLSAVLALLRAQPPCPKTKCVVDDAAQCSCGSAHIAELGLPTNLTHILLFRM 60  
Db |||||  
QY 1 MLRSALLSAVLALLRAQPPCPKTKCVVDDAAQCSCGSAHIAELGLPTNLTHILLFRM 60  
Db |||||

QY 61 DQGILRNHSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTLLRTRNKISRLPRAILD 120  
Db |||||

QY 61 DQGILRNHSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTLLRTRNKISRLPRAILD 120  
Db |||||

QY 121 MVLLLEQLFLDHNALRDLDQNLFOQLRNQLQELGLNQNLQSLFELPANLFFSSLRKLLDLSRN 180  
Db |||||

QY 121 MVLLLEQLFLDHNALRDLDQNLFOQLRNQLQELGLNQNLQSLFELPANLFFSSLRKLLDLSRN 180  
Db |||||

QY 181 NLTHLPKGLLGAQVKLEKLLLYSNQSLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDR 240  
Db |||||

QY 181 NLTHLPKGLLGAQVKLEKLLLYSNQSLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDR 240  
Db |||||

QY 241 LGNLSLTLSCGNLLESPLPALFLHVSSVSRLLTFENPLELDPDLFGEMAGRELWLNGT 300  
Db |||||

QY 241 LGNLSLTLSCGNLLESPLPALFLHVSSVSRLLTFENPLELDPDLFGEMAGRELWLNGT 300  
Db |||||

QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALR 360  
Db |||||

QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALR 360  
Db |||||

QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389  
Db |||||

QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389  
Db |||||

RESULT 2  
AAU78891  
ID AAU78891 standard; protein; 566 AA.  
XX  
AC AAU78891;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE Mouse endothelial cell specific glycoprotein GP V.  
XX  
KW Endothelial cell specific glycoprotein; GP V; anticoagulant;  
KW platelet activation inhibitor; thrombin-induced activity;  
KW transgenic animal; anti-thrombotic agent; arterial thrombosis;  
KW atherosclerotic artery; acute myocardial infarction; stroke;  
KW ischaemic complication of cardiovascular disease;  
KW Bernard-Soulier syndrome; BSS; mouse.  
XX  
OS Mus sp.  
XX  
PN WO200217711-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 31-AUG-2001; 2001WO-US026936.  
XX  
PR 31-AUG-2000; 2000US-0229047P.  
PR 31-AUG-2000; 2000US-0230566P.  
XX  
PA (CORT-) COR THERAPEUTICS INC.  
XX  
PI Ramakrishnan V, Phillips D;  
XX  
DR WPI; 2002-315494/35.

DR N-PSDB; ABK48274.  
XX  
PT Identifying agent inhibiting thrombin-induced platelet activation, by  
PT administering test agent and proteolytically inactive thrombin to GP V  
PT null non-human transgenic animal and monitoring platelet aggregation.  
XX  
PS Example 1; Fig 2; 64pp; English.  
XX  
CC The invention describes a method of identifying an agent that inhibits  
CC thrombin-induced activity (e.g., platelet activation). Thrombin-induced  
CC activity can be modulated (inhibited) by endothelial cell specific  
CC glycoprotein GP V, by administering test agent and (proteolytically  
CC inactive) thrombin to a GP V null non-human transgenic animal (I), and  
CC monitoring aggregation of platelets in (I) to identify inhibition. The  
CC method is useful for identifying agents that inhibit thrombin-induced  
CC activity e.g. thrombin-induced platelet activation, and for screening  
CC anti-thrombotic agents. Inhibitors of thrombin-induced activity are  
CC useful for inhibiting arterial thrombosis in atherosclerotic arteries  
CC e.g. acute myocardial infarction and stroke, inducing ischaemic  
CC complications of cardiovascular disease and platelet activation in a  
CC subject and may be useful for treatment of Bernard-Soulier syndrome  
CC (BSS). This is the amino acid sequence of the murine endothelial cell  
CC specific glycoprotein GP V isolated in the invention  
XX  
SQ Sequence 566 AA;

Query Match 100.0%; Score 1938; DB 5; Length 566;  
Best Local Similarity 100.0%; Pred. No. 2e-170;  
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSALLSAVLALLRAQPPCPKTKCVVDDAAQCSCGSAHIAELGLPTNLTHILLFRM 60  
Db |||||

QY 1 MLRSALLSAVLALLRAQPPCPKTKCVVDDAAQCSCGSAHIAELGLPTNLTHILLFRM 60  
Db |||||

QY 61 DQGILRNHSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTLLRTRNKISRLPRAILD 120  
Db |||||

QY 61 DQGILRNHSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTLLRTRNKISRLPRAILD 120  
Db |||||

QY 121 MVLLLEQLFLDHNALRDLDQNLFOQLRNQLQELGLNQNLQSLFELPANLFFSSLRKLLDLSRN 180  
Db |||||

QY 121 MVLLLEQLFLDHNALRDLDQNLFOQLRNQLQELGLNQNLQSLFELPANLFFSSLRKLLDLSRN 180  
Db |||||

QY 181 NLTHLPKGLLGAQVKLEKLLLYSNQSLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDR 240  
Db |||||

QY 181 NLTHLPKGLLGAQVKLEKLLLYSNQSLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDR 240  
Db |||||

QY 241 LGNLSLTLSCGNLLESPLPALFLHVSSVSRLLTFENPLELDPDLFGEMAGRELWLNGT 300  
Db |||||

QY 241 LGNLSLTLSCGNLLESPLPALFLHVSSVSRLLTFENPLELDPDLFGEMAGRELWLNGT 300  
Db |||||

QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALR 360  
Db |||||

QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALR 360  
Db |||||

QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389  
Db |||||

QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389  
Db |||||

RESULT 3  
AAR71294  
ID AAR71294 standard; protein; 560 AA.  
XX  
AC AAR71294;  
XX  
DT 25-MAR-2003 (revised)  
DT 18-AUG-1995 (first entry)  
XX  
DE Human glycoprotein V.  
XX  
KW Glycoprotein V; GPV; platelet.  
XX  
OS Homo sapiens.



```
XX SQ Sequence 560 AA;
Query Match 76.2%; Score 1477; DB 3; Length 560;
Best Local Similarity 76.1%; Pred. No. 1e-127;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 1 MRSALLSAVLALLRAQPPCPKTKCVVDDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60
DB 1 MARGTLLCAVLGLLRAQPPCPKCKVFRDAAQCSGGDVARIASALGLPTNLTHILLFGM 60
QY 61 DQILRNHSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTIRLNRKISRLEPAILDK 120
DB 61 GRGVLSQSFSFGMTVLQRLMISDSHISAVAPGTFSDLIKLTIRLNRKITHLPGALLDK 120
QY 121 MVLLEQLFLDHNAIRDLDQNLFOQLRNQLQELGNQNLSPANLFSLSRELKLLDLSRN 180
DB 121 MVLLEQLFLDHNAIRDLDQNLFOQLRNQLQELGNQNLSPANLFSLSRELKLLDLSGN 180
QY 181 NLTHLPKGLLGAQVKLEKLLLYSNQLTSDVSGLLSNLQALTELRLERNHLRSVAPGAFDR 240
DB 181 NLTHLPKGLLGAQVKLEKLLLYSNQLTSDVSGLLSNLQALTELRLERNHLRSVAPGAFDR 240
QY 241 LGNLSLTLSCNLLLESPLPALFLHVSSVSRLTLFENPLEBPDVLFGEAGLREIWLNGT 300
DB 241 LPNLSLTLSRNHLAFLPSALFLSHNLTLLTLFENPLAELPGVLFGEAGLREIWLNR 300
QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPLSALPRGVFQGLRELRLVLAHTNALAEHRDDALR 360
DB 301 QLRTLPAAAFRNLSRLRYLGVTLSPRLSALPQGAFOGLGELQVLAHNSGLTALPDGLLR 360
QY 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389
DB 361 GLGKLQVSLRNHLRALPRTLFRNLSSL 389

RESULT 5
AAU78892
ID AAU78892 standard; protein; 560 AA.
XX AC AAU78892;
XX DT 02-JUL-2002 (first entry)
XX DE Human endothelial cell specific glycoprotein GP V.
XX KW Endothelial cell specific glycoprotein; GP V; anticoagulant;
XX KW platelet activation inhibitor; thrombin-induced activity;
XX KW transgenic animal; anti-thrombotic agent; arterial thrombosis;
XX KW atherosclerotic artery; acute myocardial infarction; stroke;
XX KW ischaemic complication of cardiovascular disease;
XX KW Bernard-Soulier syndrome; BSS; human.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Misc-difference 162 /note= "Encoded by CCC"
FT Misc-difference 419 /note= "Encoded by GCG"
FT Misc-difference 485 /note= "Encoded by CAC"
FT Misc-difference 490 /note= "Encoded by GDC"
XX PN WO200217711-A2.
XX PD 07-MAR-2002.
XX PF 31-AUG-2001; 2001WO-US026936.
XX PR 31-AUG-2000; 2000US-0229047P.
XX PR 31-AUG-2000; 2000US-0230566P.
```

```
XX PA (CORT-) COR THERAPEUTICS INC.
XX PI Ramakrishnan V, Phillips D;
XX WPI; 2002-315494/35.
XX N-PSDB; ABK48275.
XX Identifying agent inhibiting thrombin-induced platelet activation, by
XX administering test agent and proteolytically inactive thrombin to GP V
XX null non-human transgenic animal and monitoring platelet aggregation.
XX Disclosure; Fig 4; 64pp; English.
XX The invention describes a method of identifying an agent that inhibits
XX thrombin-induced activity (e.g., platelet activation). Thrombin-induced
XX activity can be modulated (inhibited) by endothelial cell specific
XX glycoprotein GP V, by administering test agent and (proteolytically
XX inactive) thrombin to a GP V null non-human transgenic animal (I), and
XX monitoring aggregation of platelets in (I) to identify inhibition. The
XX method is useful for identifying agents that inhibit thrombin-induced
XX activity e.g. thrombin-induced platelet activation, and for screening
XX anti-thrombotic agents. Inhibitors of thrombin-induced activity are
XX useful for inhibiting arterial thrombosis in atherosclerotic arteries
XX e.g. acute myocardial infarction and stroke, inducing ischaemic
XX complications of cardiovascular disease and platelet activation in a
XX subject and may be useful for treatment of Bernard-Soulier syndrome
XX (BSS). This is the amino acid sequence of the human endothelial cell
XX specific glycoprotein GP V isolated in the invention
XX SQ Sequence 560 AA;
Query Match 76.2%; Score 1477; DB 5; Length 560;
Best Local Similarity 76.1%; Pred. No. 1e-127;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 1 MRSALLSAVLALLRAQPPCPKTKCVVDDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60
DB 1 MARGTLLCAVLGLLRAQPPCPKCKVFRDAAQCSGGDVARIASALGLPTNLTHILLFGM 60
QY 61 DQILRNHSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTIRLNRKISRLEPAILDK 120
DB 61 GRGVLSQSFSFGMTVLQRLMISDSHISAVAPGTFSDLIKLTIRLNRKITHLPGALLDK 120
QY 121 MVLLEQLFLDHNAIRDLDQNLFOQLRNQLQELGNQNLSPANLFSLSRELKLLDLSRN 180
DB 121 MVLLEQLFLDHNAIRDLDQNLFOQLRNQLQELGNQNLSPANLFSLSRELKLLDLSGN 180
QY 181 NLTHLPKGLLGAQVKLEKLLLYSNQLTSDVSGLLSNLQALTELRLERNHLRSVAPGAFDR 240
DB 181 NLTHLPKGLLGAQVKLEKLLLYSNQLTSDVSGLLSNLQALTELRLERNHLRSVAPGAFDR 240
QY 241 LGNLSLTLSCNLLLESPLPALFLHVSSVSRLTLFENPLEBPDVLFGEAGLREIWLNGT 300
DB 241 LPNLSLTLSRNHLAFLPSALFLSHNLTLLTLFENPLAELPGVLFGEAGLREIWLNR 300
QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPLSALPRGVFQGLRELRLVLAHTNALAEHRDDALR 360
DB 301 QLRTLPAAAFRNLSRLRYLGVTLSPRLSALPQGAFOGLGELQVLAHNSGLTALPDGLLR 360
QY 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389
DB 361 GLGKLQVSLRNHLRALPRTLFRNLSSL 389

RESULT 6
ADF69095
ID ADF69095 standard; protein; 560 AA.
XX AC ADF69095;
XX DT 12-FEB-2004 (first entry)
XX
```



DE Human MP53 protein sequence SEQ ID NO:65.  
XX p53 pathway modulating agent; MP53; p53 modulator; cytostatic;  
KW gene therapy; cancer; human.  
XX Homo sapiens.  
OS  
PN WO2003083047-A2.  
XX  
XX 09-OCT-2003.  
PD  
XX 28-FEB-2003; 2003WO-US006025.  
XX  
XX 01-MAR-2002; 2002US-0361196P.  
PR  
XX (EXEL-) EXELIXIS INC.  
PA  
XX Belvin M, Francis-Lang H, Friedman L, Plowman L, Heuer TS, Li D;  
PI Funke RP;  
PI  
XX WPI; 2003-812540/76.  
DR N-PSDB; ADF69151.  
DR  
XX Identifying a candidate p53 pathway modulating agent for treating e.g.,  
PT cancer by contacting an assay system comprising a MP53 polypeptide or  
PT nucleic acid with a test agent and detecting a test agent-biased  
PT activity.  
XX  
XX Example; SEQ ID NO 65; 406pp; English.  
XX  
XX The present invention describes a method for identifying a candidate p53  
CC pathway modulating agent, which comprises: (a) providing an assay system  
CC comprising a MP53 (modulator of p53) polypeptide or nucleic acid or its  
CC fragment or derivative; (b) contacting the assay system with a test agent  
CC under conditions where the system provides a reference activity except in  
CC the presence of the test agent; and (c) detecting a test agent-biased  
CC activity, where a difference between the test agent-biased activity and  
CC the reference activity identifies the test agent as a candidate p53  
CC pathway modulating agent. Also described: (1) modulating the p53 pathway  
CC of a cell; (2) modulating the p53 pathway in a mammalian cell; and (3)  
CC diagnosing a disease in a patient. MP53 has cytostatic activity, and can  
CC be used in gene therapy. The method is useful for identifying a candidate  
CC p53 pathway modulating agent for preparing a composition for diagnosing  
CC or treating e.g., cancer. The present sequence represents a human MP53  
CC protein, which is used in the exemplification of the present invention.  
XX  
SQ Sequence 560 AA;  
  
Query Match 76.2%; Score 1477; DB 7; Length 560;  
Best Local Similarity 76.1%; Pred. No. 1e-127;  
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;  
  
QY 1 MLRSALLSAVLALLRAQPPCPKTKCVRDAAQCSCGGVVAHIAELGLPTNLTHILLFRM 60  
Db 1 MLRGTLCAVLGLLLRAQPPCPACKCVFRDAAQCSCGGDVARISALGLPTNLTHILLFGM 60  
QY 61 DQGILRNHSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTILRLTRNKISRLPRAILDK 120  
Db 61 GRGVLOSQSFSGMTVLQRLMISDSHISAVAPGTFSDLIKLTILRSRNKITHLPGALLDK 120  
QY 121 MVLLLEQLFDHNAIRDLDQNLFOQLRNQLQELGLNQNLSPANLPSLSRELKLLDLSRN 180  
Db 121 MVLLLEQLFDHNAIRGIDQNNFQKLVNLQELALNQNLDFLPASLFTNLLENKLLDLSGN 180  
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDR 240  
Db 181 NLTHLPKGLLGAQAKLERLLHSNRLVSLDSGLLSNLGALTTELQFHRNHIRSIAPGAFDR 240  
QY 241 LGNLSSTLTSGNLLESPPALFLHVSSVSRLLTFENPLEELPDVLFEGMAGLRELWLNGT 300  
Db 241 LPNLSSTLTSRNHLAFLPSALFLHSHNLTLTLTFENPLAELPGVLFEGMGGLQELWLNR 300  
QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLALHTNALAELRDDALR 360

Db 301 QLRTLPAAAFRNLSRLRYLGVTLSPRLSALPQGAFOGLGELQVLALHSNGLTALPDGLLR 360  
QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389  
Db 361 GLGKLRQVSLRRNRLRALPRALFRNLSSL 389  
  
RESULT 7  
ABB53264  
ID ABB53264 standard; protein; 581 AA.  
XX  
AC ABB53264;  
XX  
DT 12-FEB-2002 (first entry)  
XX  
DE Human polypeptide #4.  
XX  
KW Human; nootropic; neuroprotective; anticonvulsant; antidepressant;  
KW neuroleptic; tranquiliser; antiarrhythmic; cardiant; antiasthmatic;  
KW antiinflammatory; antilipaemic; hepatotropic; virucide; antidiabetic;  
KW nephrotropic; anorectic; cytosstatic; vaccine; neurological disease;  
KW cardiovascular disease; respiratory disease; liver disease;  
KW renal disease; skeletal muscle disease; gastrointestinal disease;  
KW placental disease; testicular cancer; male fertility; pancreatic disease.  
XX  
OS Homo sapiens.  
XX  
XX WO200181363-A1.  
XX  
XX 01-NOV-2001.  
XX  
XX 26-APR-2001; 2001WO-US013360.  
XX  
XX 27-APR-2000; 2000US-0199963P.  
PR 11-MAY-2000; 2000US-0203336P.  
PR 25-MAY-2000; 2000US-0207087P.  
PR 26-MAY-2000; 2000US-0207546P.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
XX  
PI Agarwal P, Murdock PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;  
PI Lai Y, Xie Q;  
XX  
DR WPI; 2002-041392/05.  
DR N-PSDB; ABA90329.  
XX  
PS Novel polypeptides and polynucleotides useful as a vaccine for preventing  
PT and treating diseases associated the polypeptide, e.g. Alzheimer's  
PT disease, dyslipidemia, obesity, diabetes, infertility, asthma, amnesias.  
XX  
PS Claim 1; Page 70-71; 116pp; English.  
XX  
CC The invention relates to an isolated polypeptide comprising a 277, 480,  
CC 583, 581, 628, 424, 638, 229, 310, 841, 241, 369, 382, 185, 586, 1026,  
CC 844, 782, 262, 394, 471, 485, 286, 533, 495, 350, 619, 490, 462, 255,  
CC 784, 252, 593, 472, 607, 781, 640, 686 or 154 amino acid sequence as  
CC given in the specification. The polypeptides, modulators of the  
CC polypeptides and antibodies against the polypeptides are useful for  
CC treating diseases such as neurological and psychiatric diseases including  
CC Alzheimer's, parasupranuclear palsy, Huntington's disease, myotonic  
CC dystrophy, anorexia and depression; cardiovascular diseases including  
CC congestive heart failure, Hodgkin's disease and myocardial infarction;  
CC respiratory diseases including asthma, chronic obstructive pulmonary  
CC disease, cystic fibrosis and adult respiratory distress syndrome; liver  
CC diseases including hypercholesterolaemia, cirrhosis, viral and nonviral  
CC hepatitis, Type II diabetes mellitus, and impaired glucose tolerance;  
CC renal disease including renal failure, acute tubular necrosis and  
CC glomerulonephritis; skeletal muscle diseases including Eulenburg's  
CC disease, hypoglycaemia and obesity; gastrointestinal diseases including  
CC myotonia congenita and intestinal obstruction; lymph diseases including  
CC lymphagiectasia; diseases of placenta including choriocarcinoma; diseases

CC	of testes including testicular cancer, male reproductive diseases
CC	including low testosterone and male infertility; and disease of pancreas
CC	including diabetic ketoacidosis, Type 1 and 2 diabetes and obesity. The
CC	present sequence is a polypeptide of the invention
XX	Sequence 581 AA;
QY	Query Match 31.3%; Score 607; DB 5; Length 581;
Db	Best Local Similarity 37.7%; Pred. No. 4.4e-47;
QY	Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;
QY	21 CPKTKCVRDAAQCSGGSVAHIAELGLPTNLTHILLFRMDQILRNHFSGMTVLQRLM 80
Db	25 CPSECTCSRASQVECTGARIVAV-PTPLPWNAMSLQILNTHITELNESPFNLISALIALR 83
QY	81 LSDSHISAIDPGTNDLVKLTLRTRNKISRLPRAILDKMLLEQLFLDHNALRDLQDN 140
Db	84 IEKNELSRITPGAFRNLSRLYLSLANNKIQVLPIGLFGQLDSLESLLSSNQLLIQPA 143
QY	141 LFOQLRNQLQELGNQNLQSLFPLPANLFSRLRELKLLDLRSNNLTHLPKGLLGAQVKLEKLL 200
Db	144 HFSQCSNLKELQLHGNHLEYIPDGAFDHLVGLTKNLGNKSLTHISPRVFQHLGNLQVLR 203
QY	201 LYSNQLTSVDSGLLSNLGALTTELRLRNHLSVAPGAFDRLGNLSSLTLSGNLLESPPA 260
Db	204 LYENRLTDIPMGTFDGLVNLQELALQNOIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263
QY	261 LFLHVSSVSRLLTLENPLEELPDVLFGENAGLRELWLNGLHSTLTPAAAFRNLSGLQTLG 320
Db	264 IFMQLPQLNRLTLFGNSLKELSLGIFGPNRLRELWLYDNIHSSLPDNLVFSNLRQLQVLI 323
QY	321 LTRNPRLSALPRGVFQGLRELRLVLAHTNALAELRDDALRGHLRQVSLRHNRLRALPR 380
Db	324 LSRN-QISFISPGAEENGLTELRELSTLHTNALQDLDGNVFRMLANQNISLQNNRLQLPG 382
QY	381 TLFERNLSSL 389
Db	383 NIFANVNGL 391
RESULT 8	
ABU52615	
ID	ABU52615 standard; protein; 581 AA.
XX	
AC	ABU52615;
XX	
DT	04-MAR-2003 (first entry)
XX	
DE	Human NOVX protein, NOV21.
XX	
KW	Human; immunogen; NOVX; metabolic disorder; diabetes; cardiomyopathy;
KW	obesity; infectious disease; anorexia; neurodegenerative disorder;
KW	Alzheimer's disease; Parkinson's disease; immune disorder;
KW	haematopoietic disorder; dyslipidaemia; metabolic disturbance;
KW	metabolic syndrome X; wasting disorder; cancer; gene therapy; SNP;
XX	single nucleotide polymorphism.
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 11 /note= "May be Ala as a result of a single nucleotide
FT	polymorphism"
FT	Misc-difference 148 /note= "May be Arg as a result of a single nucleotide
FT	polymorphism"
FT	Misc-difference 430 /note= "May be Gly as a result of a single nucleotide
FT	polymorphism"
FT	Misc-difference 495 /note= "May be Asn as a result of a single nucleotide
FT	polymorphism"
XX	

PN	WO200281518-A2.
XX	
PD	17-OCT-2002.
XX	
PF	21-FEB-2002; 2002WO-US005374.
XX	
PR	21-FEB-2001; 2001US-0270220P.
PR	21-FEB-2001; 2001US-0270523P.
PR	23-FEB-2001; 2001US-0270797P.
PR	23-FEB-2001; 2001US-0270810P.
PR	08-MAR-2001; 2001US-0274295P.
PR	16-MAR-2001; 2001US-0276400P.
PR	16-MAR-2001; 2001US-0276677P.
PR	26-MAR-2001; 2001US-0278796P.
PR	04-APR-2001; 2001US-0281521P.
PR	25-APR-2001; 2001US-0286548P.
PR	17-MAY-2001; 2001US-0291765P.
PR	10-AUG-2001; 2001US-0311595P.
PR	13-AUG-2001; 2001US-0311980P.
PR	10-SEP-2001; 2001US-0318526P.
PR	17-SEP-2001; 2001US-0322712P.
PR	18-OCT-2001; 2001US-0330307P.
XX	
PA	(CURA-) CURAGEN CORP.
XX	
PI	Pena CEA, Shimkets RA, Li L, Shenoy SG, Kekuda R, Spytek KA;
PI	Vernet CAM, Malyankar UL, Guo X, Gusev VY, Casman SJ, Boldog FL;
PI	Furtak K, Tchernev VT, Patturajan M, Gangolli BA, Padigaru M, Liu X;
PI	Baumgartner JC, Gerlach VL, Spaderna SK, Zerhusen BD;
XX	
DR	WPI; 2003-046859/04.
DR	N-PSDB; ABX70670.
XX	
PT	New isolated NOVX polypeptide useful for treating cardiomyopathy,
PT	atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT	disease, anorexia, neurodegenerative disorders, Alzheimer's disease and
PT	cancer.
XX	
PS	Claim 1; Page 179; 479pp; English.
XX	
CC	The invention relates to an isolated polypeptide termed NOVX (NOV1, 2a,
CC	2b, 3a, 3b, 4a, 4b, 5, 6, 7a-c, 8a-e, 9a-b, 10, 11, 12a-c, 13, 14, 15,
CC	16a-d, 17a-b, 18, 19, 20a-b, 21-30) appearing as ABU52578-ABU52624), a
CC	variant of NOVX, a mature form of NOVX, and a variant of the mature form
CC	of NOVX. Also included are a nucleic acid molecule (NOVX NA) encoding
CC	NOVX, or a fragment or complement of NOVX NA, a vector comprising NOVX
CC	NA, a cell comprising the vector, an anti-NOVX antibody (ab), determining
CC	the presence or amount of NOVX or NOVX NA in a sample, and identifying an
CC	agent that binds or modulates the expression or activity of NOVX. NOVX,
CC	NOVX NA or ab is useful for treating or preventing a NOVX-associated
CC	disorder in a subject, preferably human. Ab is useful for determining the
CC	presence or amount of NOVX in a sample. NOVX is useful for identifying an
CC	agent that binds to NOVX. NOVX, NOVX NA or ab is useful for treating
CC	metabolic disorders, diabetes, cardiomyopathy, obesity, infectious
CC	disease, anorexia, neurodegenerative disorders, Alzheimer's disease,
CC	Parkinson's disease, immune disorders, haematopoietic disorders, and
CC	various dyslipidaemias, metabolic disturbances associated with obesity,
CC	the metabolic syndrome X and wasting disorders associated with chronic
CC	diseases, various cancers, endocrine, connective tissue, blood, vascular,
CC	skin, renal, bone, brain, muscle disorders, or bacterial, fungal,
CC	protozoal or viral infections. NOVX, NOVX NA or ab is useful in screening
CC	assays, detection assays, predictive medicine, and in methods of
CC	treatment. NOVX is useful as immunogen, to screen for potential
CC	ant/agonist compounds, and as bait protein in a two-hybrid or three-
CC	hybrid assay. NOVX NA is useful in gene therapy, to express NOVX, to
CC	detect NOVX mRNA or a genetic lesion in a NOVX gene, and to modulate NOVX
CC	activity. The cell is useful for producing non-human transgenic animals.
CC	Ab is useful for isolating, and purifying NOVX and to monitor protein
CC	levels in tissue as part of a clinical testing procedure. The present
CC	sequence represents a NOVX protein
XX	
SQ	Sequence 581 AA;

Query Match 31.3%; Score 607; DB 6; Length 581;  
Best Local Similarity 37.7%; Pred. No. 4.4e-47;  
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTKCVVRDAAQCSCGSAHIAELGLPTNLTHILLFRMDQGIILRNHFSFGMTVLQRLM 80  
Db 25 CPSECTCSRASQVECTGARIVAV-PTPLPWNAMSLQILNTHITELNESPFLNISALIALR 83  
QY 81 LSDSHISAIDPGTFNDLVKLTIRLNRKISRLPRAILDKMVLLEQLFLDHNALRDLDQN 140  
Db 84 IEKNELSRITPGAFRNLGSLRYLSLANNKQLVLPGLFGQLDSLESLLSSNQLLIQIPA 143  
QY 141 LFOQLRNQLQELGLNQNLSPFLPANLFSSRLRELKLLDLRSNNLTHLPKGLLGAQVKLEKLL 200  
Db 144 HFSQCSNLKELQHGHNHLEYIPDGAFDHLVGLTKNLGKNSLTHISPRVFQHLGNLQVLR 203  
QY 201 LYSNQLTSVDSGLLSNLGALTELRLERNHLRSVAPGAFDRGLGNLSSLTLSGNLLESPLPA 260  
Db 204 LYENRLTDIPMGTFDGLVNLQELALQQNQIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263  
QY 261 LFLHVSSVSRLTLFENPLEELPDVLFEGMAGLRELWNGTHLSTLPAAAFRNLSGLQTLG 320  
Db 264 IFMQLPQLNRLTLFGNSLKELSLGIFGMPNRLRELWLYDNHISSLPDNVSFNLRLQVLI 323  
QY 321 LTRNPRLSALPRGVFQGLRELRLVIALHTNALAEIRDALRGHLRQVSLRHNRLRALPR 380  
Db 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDLDCGNVFRMLANLQNLNRLRLQPLG 382  
QY 381 TLFRNLSSL 389  
Db 383 NIFANVNGL 391

RESULT 9  
ABUS6706  
ID ABUS6706 standard; protein; 581 AA.  
XX  
AC ABUS6706;  
XX  
DT 02-APR-2003 (first entry)  
XX  
DE Lung cancer-associated polypeptide #299.  
XX  
KW Lung cancer-associated polypeptide; cytostatic; emphysema;  
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
XX  
OS Unidentified.  
XX  
PN WO200286443-A2.  
XX  
PD 31-OCT-2002.  
XX  
PF 18-APR-2002; 2002WO-US012476.  
XX  
PR 18-APR-2001; 2001US-0284770P.  
PR 10-MAY-2001; 2001US-0290492P.  
PR 09-NOV-2001; 2001US-0339245P.  
PR 13-NOV-2001; 2001US-0350666P.  
PR 29-NOV-2001; 2001US-0334370P.  
PR 12-APR-2002; 2002US-0372246P.  
XX  
PA (EOSB-) EOS BIOTECHNOLOGY INC.  
XX  
PI Aziz N, Murray R;  
XX  
DR WPI; 2003-093161/08.  
DR N-PSDB; ABX76435.  
XX  
PT Detecting a lung cancer-associated transcript in a cell from a patient  
PT for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased  
PT expression in lung cancer.  
XX  
PS Claim 27; Page 423; 453pp; English.  
XX  
CC The invention relates to a method for detecting a lung cancer-associated  
CC transcript in a cell from a patient, comprising contacting a biological  
CC sample from the patient with a polynucleotide that selectively hybridises  
CC to a sequence that is at least 80 % identical to a gene that exhibits  
CC increased or decreased expression in lung cancer samples. Lung cancer-  
CC associated polynucleotides and polypeptides are used for identifying a  
CC compound that modulates a lung cancer-associated polypeptide, for  
CC inhibiting proliferation of a lung cancer-associated cell to treat lung  
CC cancer in a patient and for treating a mammal having lung cancer by  
CC administering a modulatory compound identified. The methods are useful  
CC for treating lung cancer, such as small cell lung cancer, non-small cell  
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,  
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,  
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and  
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful  
CC for diagnostic purposes and as targets for screening for therapeutic  
CC compounds that modulate lung cancer, such as antibodies. Sequences  
CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the  
CC invention  
XX  
SQ Sequence 581 AA;

Query Match 31.3%; Score 607; DB 6; Length 581;  
Best Local Similarity 37.7%; Pred. No. 4.4e-47;  
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTKCVVRDAAQCSCGSAHIAELGLPTNLTHILLFRMDQGIILRNHFSFGMTVLQRLM 80  
Db 25 CPSECTCSRASQVECTGARIVAV-PTPLPWNAMSLQILNTHITELNESPFLNISALIALR 83  
QY 81 LSDSHISAIDPGTFNDLVKLTIRLNRKISRLPRAILDKMVLLEQLFLDHNALRDLDQN 140  
Db 84 IEKNELSRITPGAFRNLGSLRYLSLANNKQLVLPGLFGQLDSLESLLSSNQLLIQIPA 143  
QY 141 LFOQLRNQLQELGLNQNLSPFLPANLFSSRLRELKLLDLRSNNLTHLPKGLLGAQVKLEKLL 200  
Db 144 HFSQCSNLKELQHGHNHLEYIPDGAFDHLVGLTKNLGKNSLTHISPRVFQHLGNLQVLR 203  
QY 201 LYSNQLTSVDSGLLSNLGALTELRLERNHLRSVAPGAFDRGLGNLSSLTLSGNLLESPLPA 260  
Db 204 LYENRLTDIPMGTFDGLVNLQELALQQNQIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263  
QY 261 LFLHVSSVSRLTLFENPLEELPDVLFEGMAGLRELWNGTHLSTLPAAAFRNLSGLQTLG 320  
Db 264 IFMQLPQLNRLTLFGNSLKELSLGIFGMPNRLRELWLYDNHISSLPDNVSFNLRLQVLI 323  
QY 321 LTRNPRLSALPRGVFQGLRELRLVIALHTNALAEIRDALRGHLRQVSLRHNRLRALPR 380  
Db 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDLDCGNVFRMLANLQNLNRLRLQPLG 382  
QY 381 TLFRNLSSL 389  
Db 383 NIFANVNGL 391

RESULT 10  
ABUS6544  
ID ABUS6544 standard; protein; 581 AA.  
XX  
AC ABUS6544;  
XX  
DT 02-APR-2003 (first entry)  
XX  
DE Lung cancer-associated polypeptide #137.  
XX  
KW Lung cancer-associated polypeptide; cytostatic; emphysema;  
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;







CC proteins given in ABR43169 to ABR43204 and designated REMAP-1 to REMAP-36  
CC (I). (I) have cytostatic, antiarteriosclerotic, anticonvulsant,  
CC nootropic, neuroprotective, cerebroprotective, anti-HIV, antiallergic,  
CC antiinflammatory and thyromimetic activities, and can be used in gene  
CC therapy. The REMAP polypeptides and polynucleotides are useful in  
CC diagnosing, treating and preventing diseases or conditions associated  
CC with the decreased expression or overexpression of REMAP, such as cell  
CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g.  
CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,  
CC allergies) and developmental (e.g. hypothyroidism, Cushing's syndrome)  
CC disorders, or infections. They are also useful in assessing the effects  
CC of exogenous compounds on the expression of nucleic acid and amino acid  
CC sequences of REMAP. The REMAPs or their fragments are useful in screening  
CC compounds for effectiveness as agonist or antagonist of the polypeptides,  
CC or in altering the expression of the target polynucleotide and compounds  
CC that specifically bind to or modulate the activity of the polypeptide

XX Sequence 581 AA;

Query Match 31.3%; Score 607; DB 6; Length 581;  
Best Local Similarity 37.7%; Pred. No. 4.4e-47;  
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKCKCVVDDAAQCSCGSAHIAELGLPTNLTHILLFRMDQGILRNHFSFGMTVLQRLM 80  
Db 25 CPSECTCSRASQVECTGARIVAV-PTPLPWAMSLQILNTHITELNESPFNLISALIALR 83  
QY 81 LSDSHISAIDPGTFNDLVKLTLLTRNKISRLPRAILDKMVLLQFLDHNALRLDQDN 140  
Db 84 IEKNELSRITPGAFRNLGSLRYLSLANNKQLVPLGFLQGLDSLESLLSSNQLLIQIPA 143  
QY 141 LFNQQLRNQLQELGNQNLSPFLPANLFSSSLRELKLLDLNRNNLTHLPKGLLGAQVKLEKLL 200  
Db 144 HFSQCSNLKELQLHGNHLEYIPDGAFDHLVGLTKNLGKNSLTHISPRVFQHLGNLQVLR 203  
QY 201 LYSNQLTSVDSGLLSNLGALTELRLERNHLRSVAPGAFDRGLGNLSSLTSGNLLSPLPA 260  
Db 204 LYENRLTDIPMGTFDGLVNLQELALQQNQIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263  
QY 261 LFLHVSSVSRLLTLFENPLEELPDVLFEGMAGRELWLNGLTHLSTLPAAAFRNLSGLQTLG 320  
Db 264 IFMQLPQLNRLTLFGNSLKELSLGIFGPMPLRELWLYDYNHISSLPDNVSFNLRLQQLVLI 323  
QY 321 LTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDDALRGLHRLQVSLRHNRLRALPR 380  
Db 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDLGDNVFRMLANLQISLQNNRLRLQLPG 382  
QY 381 TLFRLNSSL 389  
Db 383 NIFANVNGL 391

RESULT 12  
ADC83709  
ID ADC83709 standard; protein; 581 AA.

XX ADC83709;

DT 01-JAN-2004 (first entry)

DE Human leucine rich repeat (LRR) Lib protein.

XX Human; Lib; chromosome 3q29; leucine-rich repeat; LRR;  
KW transmembrane domain; adhesion; cell-cell; cell-extracellular matrix;  
KW astroglial cell; beta-amyloid processing; Alzheimer's disease;  
KW neuron death; drug screening; cell adhesion-related disease; nootropic;  
KW neuroprotective; gene therapy.

OS Homo sapiens.

PN JP2003164290-A.

XX 10-JUN-2003.

XX 30-NOV-2001; 2001JP-00367093.  
XX 30-NOV-2001; 2001JP-00367093.  
XX (DAUC ) DAIICHI PHARM CO LTD.  
XX WPI; 2003-630582/60.  
XX N-PSDB; ADC83707.  
XX Novel gene encoding a protein having cell-cell adhesion or cell-  
XX extracellular matrix adhesion regulatory property useful for diagnosing  
XX or treating neuron death, and Alzheimer's disease.  
XX Claim 3; Page 14-16; 20pp; Japanese.

XX The invention relates to rat and human proteins designated Lib (ADC83708-  
XX ADC83709), and cDNAs encoding them (ADC83706-ADC83707). The Lib proteins  
XX contain a leucine-rich repeat (LRR) and a transmembrane domain, and are  
XX involved in regulation of cell-cell adhesion or cell-extracellular matrix  
XX adhesion. They are expressed in astroglial cells and their expression is  
XX induced by beta-amyloid processing, indicating that they may be useful in  
XX the diagnosis, treatment or prevention of Alzheimer's disease. The  
XX invention also encompasses polynucleotides with at least 70% homology to  
XX the Lib polynucleotides, recombinant vectors and host cells comprising a  
XX Lib polynucleotide, the recombinant expression of Lib proteins, an  
XX antibody against a Lib protein, and methods of screening for modulators  
XX of cell adhesion. Lib proteins and their encoding nucleic acids may be  
XX used in the diagnosis of Alzheimer's disease or neuron death, and also to  
XX screen for modulators of its activity. Such compounds are useful for  
XX treating or preventing Alzheimer's disease-related neuron death or a  
XX disease involving cell-cell adhesion or cell-extracellular matrix  
XX adhesion. The present sequence represents human Lib protein. The human  
XX Lib gene encoding it is located on chromosome 3q29.

SQ Sequence 581 AA;

Query Match 31.3%; Score 607; DB 7; Length 581;  
Best Local Similarity 37.7%; Pred. No. 4.4e-47;  
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKCKCVVDDAAQCSCGSAHIAELGLPTNLTHILLFRMDQGILRNHFSFGMTVLQRLM 80  
Db 25 CPSECTCSRASQVECTGARIVAV-PTPLPWAMSLQILNTHITELNESPFNLISALIALR 83  
QY 81 LSDSHISAIDPGTFNDLVKLTLLTRNKISRLPRAILDKMVLLQFLDHNALRLDQDN 140  
Db 84 IEKNELSRITPGAFRNLGSLRYLSLANNKQLVPLGFLQGLDSLESLLSSNQLLIQIPA 143  
QY 141 LFNQQLRNQLQELGNQNLSPFLPANLFSSSLRELKLLDLNRNNLTHLPKGLLGAQVKLEKLL 200  
Db 144 HFSQCSNLKELQLHGNHLEYIPDGAFDHLVGLTKNLGKNSLTHISPRVFQHLGNLQVLR 203  
QY 201 LYSNQLTSVDSGLLSNLGALTELRLERNHLRSVAPGAFDRGLGNLSSLTSGNLLSPLPA 260  
Db 204 LYENRLTDIPMGTFDGLVNLQELALQQNQIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263  
QY 261 LFLHVSSVSRLLTLFENPLEELPDVLFEGMAGRELWLNGLTHLSTLPAAAFRNLSGLQTLG 320  
Db 264 IFMQLPQLNRLTLFGNSLKELSLGIFGPMPLRELWLYDYNHISSLPDNVSFNLRLQQLVLI 323  
QY 321 LTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDDALRGLHRLQVSLRHNRLRALPR 380  
Db 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDLGDNVFRMLANLQISLQNNRLRLQLPG 382  
QY 381 TLFRLNSSL 389  
Db 383 NIFANVNGL 391

RESULT 13  
ADE08191  
ID ADE08191 standard; protein; 581 AA.



CC whose expression is upregulated or downregulated in specific cancers or  
CC other diseases such as angiogenic or fibrotic disorders, and to methods  
CC of determining the presence or absence of a pathological cell in a  
CC patient by detecting a nucleic acid at least 80% identical to those of  
CC the invention or by detecting a polypeptide of the invention. The  
CC invention also relates to expression vectors and host cells comprising a  
CC nucleic acid of the invention; antibodies which specifically bind a  
CC polypeptide of the invention; use of such antibodies for drug targeting;  
CC and methods of screening for modulators of activity or expression of the  
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,  
CC antibodies and methods are useful for diagnosing, prognosing and treating  
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,  
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal  
CC neovascularisation syndromes, scarring and uterine fibroids. They may  
CC also be useful in wound healing and in contraception. The present  
CC sequence represents a polypeptide of the invention.

XX Sequence 581 AA;

Query Match 31.3%; Score 607; DB 7; Length 581;  
Best Local Similarity 37.7%; Pred. No. 4.4e-47;  
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTKCVVRDAAQCSCGSAHIAELGLPTNLTHILLFRMDQGILRNHFSFGMTVLQRLM 80  
DB 25 CPSECTCSRASQVECTGARIVAV-PTPLPWNAISLIQILNTHITELNESPFNLISALIALR 83  
QY 81 LSDSHISAIDPGTFNDLVKLTLRNKRISRLPRAILDKMVLLEQLFLDHNALRDLQDN 140  
DB 84 IEKNELSRITPGAFRNLGSLRYLSLANNKLQVLPFGQLDLSLESLLSSNQLLIQIPA 143  
QY 141 LFQQLRNQLQELGLNQNLISFLPANLFSLSRELKLLDLSRNNLTHLPKGLLGAQVVKLEKLL 200  
DB 144 HFSQCSNLKELQLHGNHLEYIPDGAFDHLVGLTKNLGKNSLTHISPRVFOHLGNLQVLR 203  
QY 201 LYSNQLTSVDSGLLSNLGALTELRNHLRSVAPGAPDRGLGNLSLTLTSGNLLSLPPA 260  
DB 204 LYENRLTDIPMGTFDGLVNLQELALQONQIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263  
QY 261 LFLHVSVSRLLTFENPLELPDVLFGEMAGLRELWNGTHLSTLPAAAFRNLSGLQTLG 320  
DB 264 IFMQLPQLNRLTLFGNSLKELSLGIFGPMNRLRELMWYDYNHISLFPDVFNSRLQQLVLI 323  
QY 321 LTRNPRLSALPRGVFQGLRELRLVLAHTNALAEELDDALRGLHLRQVSLRHNRRLALPR 380  
DB 324 LSRN-QISFISPGAFNGLTRELRLSLHTNALQDLQDGNVFRMLANLQNISLQNRLRLQPLG 382  
QY 381 TLPRNLSL 389  
DB 383 NIPANVNGL 391

RESULT 15  
ADN39092  
ID ADN39092 standard; protein; 581 AA.  
XX  
AC ADN39092;  
XX

17-JUN-2004 (first entry)  
Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:410.

KW Human; differential expression; cancer; angiogenic disorder;  
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
KW inflammatory disease; autoimmune disease;  
KW retinal neovascularisation syndrome; scarring; uterine fibroid;  
KW detection; diagnosis; prognosis; drug screening; drug targeting;  
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;  
KW vulnery; gene therapy; vaccine.

OS Homo sapiens.  
XX  
PN WO2003042661-A2.

XX 22-MAY-2003.  
PD  
XX 13-NOV-2002; 2002WO-US036810.  
XX  
PR 13-NOV-2001; 2001US-0350666P.  
PR 21-NOV-2001; 2001US-0332464P.  
PR 29-NOV-2001; 2001US-0334393P.  
PR 03-DEC-2001; 2001US-0335394P.  
PR 14-DEC-2001; 2001US-0340376P.  
PR 08-JAN-2002; 2002US-0347211P.  
PR 10-JAN-2002; 2002US-0347349P.  
PR 08-FEB-2002; 2002US-0355250P.  
PR 13-FEB-2002; 2002US-0356714P.  
PR 20-FEB-2002; 2002US-0359077P.  
PR 29-MAR-2002; 2002US-036809P.  
PR 04-APR-2002; 2002US-0370110P.  
PR 12-APR-2002; 2002US-0372246P.  
PR 05-JUN-2002; 2002US-0386614P.  
PR 16-JUL-2002; 2002US-0396839P.  
PR 22-JUL-2002; 2002US-0397775P.  
PR 22-JUL-2002; 2002US-0397845P.  
PR 09-SEP-2002; 2002US-0409450P.  
XX  
(EOSB-) EOS BIOTECHNOLOGY INC.

PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;  
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;  
XX  
DR WPI; 2003-468649/44.  
DR N-PSDB; ADN39091.

Determining the presence or absence of a pathological cell in a patient,  
useful for diagnosing, prognosing or treating cancer, comprises detecting  
a nucleic acid in a biological sample.

Claim 12; SEQ ID NO 410; 1385pp; English.

The invention relates to nucleic acids and proteins (ADN38683-ADN40064)  
whose expression is upregulated or downregulated in specific cancers or  
other diseases such as angiogenic or fibrotic disorders, and to methods  
of determining the presence or absence of a pathological cell in a  
patient by detecting a nucleic acid at least 80% identical to those of  
the invention or by detecting a polypeptide of the invention. The  
invention also relates to expression vectors and host cells comprising a  
nucleic acid of the invention; antibodies which specifically bind a  
polypeptide of the invention; use of such antibodies for drug targeting;  
and methods of screening for modulators of activity or expression of the  
polypeptides and nucleic acids. The nucleic acids, polypeptides,  
antibodies and methods are useful for diagnosing, prognosing and treating  
cancer and other conditions such as psoriasis, ischaemia, heart disease,  
atherosclerosis, inflammatory diseases, autoimmune diseases, retinal  
neovascularisation syndromes, scarring and uterine fibroids. They may  
also be useful in wound healing and in contraception. The present  
sequence represents a polypeptide of the invention.

XX Sequence 581 AA;

Query Match 31.3%; Score 607; DB 7; Length 581;  
Best Local Similarity 37.7%; Pred. No. 4.4e-47;  
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTKCVVRDAAQCSCGSAHIAELGLPTNLTHILLFRMDQGILRNHFSFGMTVLQRLM 80  
DB 25 CPSECTCSRASQVECTGARIVAV-PTPLPWNAISLIQILNTHITELNESPFNLISALIALR 83  
QY 81 LSDSHISAIDPGTFNDLVKLTLRNKRISRLPRAILDKMVLLEQLFLDHNALRDLQDN 140  
DB 84 IEKNELSRITPGAFRNLGSLRYLSLANNKLQVLPFGQLDLSLESLLSSNQLLIQIPA 143  
QY 141 LFQQLRNQLQELGLNQNLISFLPANLFSLSRELKLLDLSRNNLTHLPKGLLGAQVVKLEKLL 200  
DB 144 HFSQCSNLKELQLHGNHLEYIPDGAFDHLVGLTKNLGKNSLTHISPRVFOHLGNLQVLR 203

Qy	201	LYSNQTSVDSGLLSNLGALTTELRLERNHRSVAPGAFDRGLNLSLTLSCNLLLESPPA	260
Db	204	LYENRLTDIPMGTFDGLVNLQELALQQNOIGLLSPGLFHHNNHNLQRLYLSNNHISQLPPS	263
Qy	261	LFLHVSSVSRLLTFENPLEELPDVLFGEWAGIRELWNGTHLSTLPAAAFRNLSGLQTLG	320
Db	264	IFMQLPQLNRLTLFGNSLKELSLGIFGPMENIRELWLYDNHISLDPDNVFSNLRQLQVLI	323
Qy	321	LTRNPRLSALPRGVFOGLRELRVLAHHTNALAELRDDALRGHLRQVSLRHNRLRALPR	380
Db	324	LSRN-QISFISPGAENGLTELRELSLHTNALQDLDGNVFRMLANLQNISLQNNRLRQLPG	382
Qy	381	TLFRNLSSL	389
Db	383	NIFANVNGL	391

Search completed: November 29, 2004, 13:09:04  
Job time : 89 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 13:11:04 ; Search time 93 Seconds  
(without alignments)  
1483.616 Million cell updates/sec

Title: US-09-775-803-12\_COPY\_1\_389  
Perfect score: 1938  
Sequence: 1 MLRSALLSAVLALLRAQFPF.....LRHNRRLRALPRTLFRNLSSL 389

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1938	100.0	566	10	US-09-775-803-12
2	1938	100.0	567	15	US-10-037-417-105
3	1923	99.2	567	15	US-10-037-417-106
4	1749	90.2	567	15	US-10-080-334-228
5	1477	76.2	560	9	US-09-782-980-61
6	1477	76.2	560	10	US-09-775-803-14
7	1477	76.2	560	14	US-10-212-499-2
8	1477	76.2	560	16	US-10-806-018-61
9	1201	62.0	368	15	US-10-312-311-21
10	607	31.3	581	14	US-10-295-027-410
11	607	31.3	581	14	US-10-295-027-1274
12	607	31.3	581	15	US-10-080-334-76
13	607	31.3	581	15	US-10-080-334-225
					Sequence 12, Appl
					Sequence 105, App
					Sequence 106, App
					Sequence 228, App
					Sequence 61, Appl
					Sequence 14, Appl
					Sequence 2, Appli
					Sequence 61, Appl
					Sequence 21, Appl
					Sequence 410, App
					Sequence 1274, Ap
					Sequence 76, Appl
					Sequence 225, App

14	607	31.3	581	15	US-10-258-951-43	Sequence 43, Appl
15	607	31.3	787	15	US-10-080-334-224	Sequence 224, App
16	607	31.3	790	14	US-10-017-161-1558	Sequence 1558, Ap
17	607	31.3	790	14	US-10-292-798-1246	Sequence 1246, Ap
18	607	31.3	819	14	US-10-241-220-106	Sequence 106, App
19	607	31.3	819	17	US-10-872-972-106	Sequence 106, App
20	583	30.1	578	15	US-10-080-334-226	Sequence 226, App
21	541.5	27.9	391	15	US-10-080-334-227	Sequence 227, App
22	503.5	26.0	570	14	US-10-424-233-14	Sequence 14, Appl
23	445	23.0	545	15	US-10-114-270-38	Sequence 38, Appl
24	442	22.8	545	15	US-10-114-270-40	Sequence 40, Appl
25	419.5	21.6	457	15	US-10-312-311-22	Sequence 22, Appl
26	419.5	21.6	605	16	US-10-408-765A-632	Sequence 632, App
27	416.5	21.5	605	16	US-10-706-791-21	Sequence 21, Appl
28	413.5	21.3	510	15	US-10-312-311-20	Sequence 20, Appl
29	413.5	21.3	605	9	US-09-782-980-62	Sequence 62, Appl
30	413.5	21.3	605	16	US-10-806-018-62	Sequence 62, Appl
31	379.5	19.6	928	16	US-10-442-658-28	Sequence 28, Appl
32	379.5	19.6	1122	16	US-10-442-658-27	Sequence 27, Appl
33	379.5	19.6	1461	16	US-10-442-658-7	Sequence 7, Appli
34	379.5	19.6	1531	10	US-09-970-944-31	Sequence 31, Appl
35	379.5	19.6	1534	10	US-09-970-944-30	Sequence 30, Appl
36	379.5	19.6	1534	16	US-10-442-658-6	Sequence 6, Appli
37	375.5	19.4	1531	10	US-09-970-944-28	Sequence 28, Appl
38	369.5	19.1	796	14	US-10-028-392-5	Sequence 5, Appli
39	369.5	19.1	1529	10	US-09-866-050A-396	Sequence 396, App
40	367.5	19.0	907	14	US-10-271-078-10	Sequence 10, Appl
41	366.5	18.9	1348	15	US-10-080-334-234	Sequence 234, App
42	366	18.9	467	15	US-10-343-348-36	Sequence 36, Appl
43	362	18.7	421	15	US-10-072-012-294	Sequence 294, App
44	362	18.7	1523	10	US-09-954-342-45	Sequence 45, Appl
45	361.5	18.7	1529	10	US-09-766-511B-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1  
US-09-775-803-12  
; Sequence 12, Application US/09775803  
; Publication No. US20030167487A1  
; GENERAL INFORMATION:  
; APPLICANT: COR Therapeutics, Inc.  
; APPLICANT: Ramakrishnan, Vanitha  
; APPLICANT: Phillips, David  
; TITLE OF INVENTION: Transgenic Animals Having a Modified Glycoprotein V  
; TITLE OF INVENTION: Gene and Methods for Their Use  
; FILE REFERENCE: 44481-5044-US  
; CURRENT APPLICATION NUMBER: US/09/775,803  
; CURRENT FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: PCT/US99/17594  
; PRIOR FILING DATE: 1999-08-04  
; PRIOR APPLICATION NUMBER: US 60/109,797  
; PRIOR FILING DATE: 1998-08-04  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 566  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-775-803-12

Query Match 100.0%; Score 1938; DB 10; Length 566;  
Best Local Similarity 100.0%; Pred. No. 4.6e-156;  
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MLRSALLSAVLALLRAQFPFCKTKCVVRDAACQSGGSAHIAELGLPTNLTHILFRM	60
Db	1	MLRSALLSAVLALLRAQFPFCKTKCVVRDAACQSGGSAHIAELGLPTNLTHILFRM	60
Qy	61	DOGILRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKISRLPRAILDK	120
Db	61	DOGILRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKISRLPRAILDK	120

QY 121 MVLLEQLFLDHNALRDLDQNLFOQLRNLOELGLNQNLQSLFPLANLFSSRLRELKLLDLSRN 180  
Db 121 MVLLEQLFLDHNALRDLDQNLFOQLRNLOELGLNQNLQSLFPLANLFSSRLRELKLLDLSRN 180  
QY 181 NLTHLPKGLLGAQVKLEKLLLYSNQLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDR 240  
Db 181 NLTHLPKGLLGAQVKLEKLLLYSNQLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDR 240  
QY 241 LGNLSLTLTSGNLLLESPLPALFLHVSSVSRLTLFENPLEELPDVLFGEWAGLRELWLNGT 300  
Db 241 LGNLSLTLTSGNLLLESPLPALFLHVSSVSRLTLFENPLEELPDVLFGEWAGLRELWLNGT 300  
QY 301 HSLTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAELRDDALR 360  
Db 301 HSLTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAELRDDALR 360  
QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389  
Db 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389

RESULT 2

US-10-037-417-105  
; Sequence 105, Application US/10037417  
; Publication No. US20040052806A1

GENERAL INFORMATION:

; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Grosse, William M  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Li, Li  
; APPLICANT: Gorman, Linda  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: Sciore, Paul  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Rothenberg, Mark  
; APPLICANT: Stone, David J  
; APPLICANT: Boldog, Ferenc L  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Anderson, David W  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Miller, Charles E  
; APPLICANT: Eisen, Andrew J

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-235

; CURRENT APPLICATION NUMBER: US/10/037,417

; CURRENT FILING DATE: 2002-09-20

; PRIOR APPLICATION NUMBER: 60/260,018

; PRIOR FILING DATE: 2001-01-05

; PRIOR APPLICATION NUMBER: 60/260,360

; PRIOR FILING DATE: 2001-01-08

; PRIOR APPLICATION NUMBER: 60/272,411

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: 60/272,817

; PRIOR FILING DATE: 2001-03-02

; PRIOR APPLICATION NUMBER: 60/291,186

; PRIOR FILING DATE: 2001-05-15

; PRIOR APPLICATION NUMBER: 60/303,231

; PRIOR FILING DATE: 2001-07-05

; PRIOR APPLICATION NUMBER: 60/305,060

; PRIOR FILING DATE: 2001-07-12

; PRIOR APPLICATION NUMBER: 60/318,405

; PRIOR FILING DATE: 2001-09-10

; PRIOR APPLICATION NUMBER: 60/318,700  
; PRIOR FILING DATE: 2001-09-12

; NUMBER OF SEQ ID NOS: 227

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 105

; LENGTH: 567

; TYPE: PRT

; ORGANISM: Mus musculus

US-10-037-417-105

Query Match 100.0%; Score 1938; DB 15; Length 567;

Best Local Similarity 100.0%; Pred. No. 4.6e-156;

Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSALLSAVLALRAQPPPCPKTKCVVRDAAQCGSVAHIAELGLPTNLTHILLFRM 60

Db 1 MLRSALLSAVLALRAQPPPCPKTKCVVRDAAQCGSVAHIAELGLPTNLTHILLFRM 60

QY 61 DQILRNHSFGMTVLQRLMLSDSHISALDPGTFTNDLVKLKTLRLTRNKISRLPAILDK 120

Db 61 DQILRNHSFGMTVLQRLMLSDSHISALDPGTFTNDLVKLKTLRLTRNKISRLPAILDK 120

QY 121 MVLLEQLFLDHNALRDLDQNLFOQLRNLOELGLNQNLQSLFPLANLFSSRLRELKLLDLSRN 180

Db 121 MVLLEQLFLDHNALRDLDQNLFOQLRNLOELGLNQNLQSLFPLANLFSSRLRELKLLDLSRN 180

QY 181 NLTHLPKGLLGAQVKLEKLLLYSNQLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDR 240

Db 181 NLTHLPKGLLGAQVKLEKLLLYSNQLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDR 240

QY 241 LGNLSLTLTSGNLLLESPLPALFLHVSSVSRLTLFENPLEELPDVLFGEWAGLRELWLNGT 300

Db 241 LGNLSLTLTSGNLLLESPLPALFLHVSSVSRLTLFENPLEELPDVLFGEWAGLRELWLNGT 300

QY 301 HSLTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAELRDDALR 360

Db 301 HSLTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAELRDDALR 360

QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389

Db 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389

RESULT 3

US-10-037-417-106

; Sequence 106, Application US/10037417

; Publication No. US20040052806A1

GENERAL INFORMATION:

; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Grosse, William M  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Li, Li  
; APPLICANT: Gorman, Linda  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: Sciore, Paul  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Rothenberg, Mark  
; APPLICANT: Stone, David J  
; APPLICANT: Boldog, Ferenc L  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Anderson, David W  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Miller, Charles E  
; APPLICANT: Eisen, Andrew J

```

; APPLICANT: Eisen, Andrew J
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-235
; CURRENT APPLICATION NUMBER: US/10/037,417
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/260,018
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,360
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/272,411
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/272,817
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/291,186
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/303,231
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/305,060
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/318,405
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/318,700
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 227
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 106
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-037-417-106

```

Query Match	99.2%	Score 1923;	DB 15;	Length 567;
Best Local Similarity	99.2%	Pred. NO. 8.7e-155;		
Matches 386; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	1	MLRSALLSAVLALLRAQFPFPCPKTKCWVRDAAQCSCGSAHIAELGLPTNLTHILLFRM	60
Db	1	MLRSALLSAVPLPLRAQFPFPCPKTKCWVRDAAQCSCGSAHIAELGLPTNLTHILLFRM	60
QY	61	DQGI LRNHSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLT LRLTRNKISR LPRAILDK	120
Db	61	DQGI LRNHSFSGMTVLQ RQMLSDSHISAIDPGTFNDLVKLT LRLTRNKISR LPRAILDK	120
QY	121	MVLEQLFLDHNALRDLQNLFOQLRNQLQELGLNQNLQLSFLPANLFSSLR ELKL DLSRN	180
Db	121	MVLEQLFLDHNALRDLQNLFOQLRNQLQELGLNQNLQLSFLPANLFSSLR ELKL DLSRN	180
QY	181	NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELR LERNHLRSVAPGAFDR	240
Db	181	NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELR LERNHLRSVAPGAFDR	240
QY	241	LGNLSSLT LSGNLLLES LPPALFLHVSSVSRLTLFENPIUELPDVLFGEMAGRLWLNGT	300
Db	241	LGNLSSLT LSGNLLLES LPPALFLHVSSVSRLTLFENPIUELPDVLFGEMAGRLWLNGT	300
QY	301	HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGI RELRVLALHTNALAE LRDDALR	360
Db	301	HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGI RELRVLGLHTNALAE LRDDALR	360
QY	361	GLGHLRQVSLRHNRLRALPRTLFRNLSSL	389
Db	361	GLGHLRQVSLRHNRLRALPRTLFRNLSSL	389

RESULT 4  
US-10-080-334-228  
; Sequence 228, Application US/10080334  
; Publication No. US20040002584A1  
; GENERAL INFORMATION:  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Li, Li  
; APPLICANT: Shenoy, Suresh G

```

; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Liu, Xiaochong
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spaderna, Steven K
; APPLICANT: Zerhusen, Bryan D
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-275
; CURRENT APPLICATION NUMBER: US/10/080,334
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,523
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/322,712
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/311,980
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/330,307
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/278,796
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/281,521
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/276,677
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,595
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/270,220
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/274,295
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/318,526
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/286,548
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/291,765
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/270,797
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/276,400
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/270,810
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 228
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-080-334-228

```

```

Query Match          90.2%; Score 1749; DB 15; Length 567;
Best Local Similarity 89.2%; Pred. NO. 5.5e-140;
Matches 347; Conservative 24; Mismatches 18; Indels 0; Gaps 0;

Qy      1 MLRSALLSAVLALLRAQPPCPKCTCKCVVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60
      |||||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1 MLRSVLLSAVLSLVGAQPPFPCKTCKCVVRDAVQCSGGSVAHIAELGLPTNLTHILLFRM 60
      |||||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy     61 DQGIILRNHSPSGMTVLQRLMLSDSHISAIDPGTFNDLVKLKTLRLTRNKISRLPRAILDK 120
      ||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

Db	61	DRGVLSHSFGMTVQLRMLMSDSHISAIDPGTFTNDLVKLTRLRLTRNKISHLPRAILDK	120
QY	121	MVLEQELFDHNAIRDLDQNLFFQQLRNQLQELGLNQNLQSLFDPANLFFSSLRELKLLDLSRN	180
Db	121	MVLEQELFDHNAIRDLDQNLFFQQLRNLRDCLNQNLQSLFDPANLFFSSLGKLVKVLDSRN	180
QY	181	NLTHLPKGLHGAQVKLEKLLYSNQLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDR	240
Db	181	NLTHLPQGLHGAQIKLEKLLYSNRLMSLDSGLLANLGALTTELRLERNHLRSIAPGAFDS	240
QY	241	LGNLSLTLSGNLLLESPPALFLHVSSVSRLTLTFENPLEELPDVLFGEMAGLRELWINGT	300
Db	241	LGNLSLTLSGNLLLESPPALFLHVSWSLTRLTLTFENPLEELPEVLFGEMAGLRELWINGT	300
QY	301	HLSTLPAAAFRNLSSGLQTLGLTRNPRLSALPRGVFQGLRELKVLALHTNALAELRDDALR	360
Db	301	HLRTLPAAAFRLNSGLQTLGLTRNPLLSALPPGMFHGLTELKVLAVHTNALEELPEDALR	360
QY	361	GLGHLRQVSLRHNRLRALPRTLFRNLSSL	389
Db	361	GLGRLROVSLRHNRLRALPRTLFRNLSSL	389

## RESULTS

US-09-782-980-61  
; Sequence 61, Application US/09782980  
; Patent No. US20020072089A1  
; GENERAL INFORMATION:  
; APPLICANT: Khodadoust, Mehran M.  
; APPLICANT: MacBeth, Kyle J.  
; APPLICANT: Busfield, Samantha J.  
; APPLICANT: McCarthy, Sean A.  
; APPLICANT: Holtzman, Douglas A.  
; APPLICANT: Gu, Wei  
; APPLICANT: White, David  
; APPLICANT: Pan, Yang  
; TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIFE, TRASH, LRSG, AND  
; TITLE OF INVENTION: STMST PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: MNI-121CP  
; CURRENT APPLICATION NUMBER: US/09/782,980  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: PCT/US00/02125  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: 09/448,076  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: 09/276,400  
; PRIOR FILING DATE: 1999-03-25  
; PRIOR APPLICATION NUMBER: 60/117,580  
; PRIOR FILING DATE: 1999-01-27  
; PRIOR APPLICATION NUMBER: 09/014,195  
; PRIOR FILING DATE: 1998-01-27  
; PRIOR APPLICATION NUMBER: 09/014,348  
; PRIOR FILING DATE: 1998-01-27  
; PRIOR APPLICATION NUMBER: 09/086,892  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 09/296,208  
; PRIOR FILING DATE: 1999-04-21  
; PRIOR APPLICATION NUMBER: 09/063,950  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 09/561,381  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 09/561,810  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 09/087,121  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 09/672,721  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: 09/049,799  
; PRIOR FILING DATE: 1998-03-27  
; NUMBER OF SEQ ID NOS: 176  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 61

```
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-980-61

Query Match      76.2%; Score 1477; DB 9; Length 560;
Best Local Similarity 76.1%; Pred. No. 7.4e-117;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY   1 MLRSALLSAVLALLRAQPFPCKPTCKCVVRDAAQCSCGGVAHIAELGLPTNLTHILLFRM 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1 MLRGTLCAVGLGLLRAQPFPCCPACKCVFEDAAQCSCGGDVARIASALGLPTNLTHILLFGM 60

QY   61 DQGLRNHSFSGMTVLQRMLSDSHISAIDPGTFNDLVKLKTLRLTRNKISRLPRLDK 120
    :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db   61 GRGYLQSQSFSGMTVLQRLMISDSHISAVAPGTFSDLIKLTKLRLSRINKITHLPGLLDK 120

QY   121 MYLLEQLFLDHNALRDLDQNLFQQLRNLQELGLNQNLSTFLPANLFFSSLRELKLLDLNRN 180
    ||| ||| ||| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db   121 MYLLEQLFLDHNALRGIDQNMFMFKLVNLQELALNQNLDFLPASLTNLENLKLLDLNSN 180

QY   181 NLTHLPKGLLGAAQVKLEKLLLYSNQLTSVDSGLLSLNGALTTELRLERNHLRSVAPGAFDR 240
    ||| ||| ||| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db   181 NLTHLPKGLLGAAQAKLERLLHNSRNLVSILDSGLLSLNGALTTELOFHRNHIRSIAPGAFDR 240

QY   241 LGNLSLTLTSGNLESPLPALFLHVSSVSRLTLFENPLEHPDVLPGEMAGLREIWLNGT 300
    ||| ||| ||| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db   241 LFNLSLTLTSGNLESPLPALFLHSHNLTLTLFENPLAELPGVLFGEMGGLQELWLNRT 300

QY   301 HUSTLPAAAFRNLSGLIQTGLTRNPRLSALPRGVFQGLRELRVLALHTNALAEIRDALR 360
    ||| ||| ||| ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db   301 QLRTLPAFAFRNLSRYLGVTLSPRLSALPDQGAFOGIGELQVLALHNSGLTALPDGLLR 360

QY   361 GLCHLRQVSLRHNRLRALPRTLFRNLSSL 389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   361 GLCKLRQVSLRRNRLRALPRALFRNLSSL 389


RESULT 6
US-09-775-803-14
; Sequence 14, Application US/09775803
; Publication No. US20030167487A1
; GENERAL INFORMATION:
; APPLICANT: COR Therapeutics, Inc.
; APPLICANT: Ramakrishnan, Vanitha
; APPLICANT: Phillips, David
; TITLE OF INVENTION: Transgenic Animals Having a Modified Glycoprotein V
; TITLE OF INVENTION: Gene and Methods for Their Use
; FILE REFERENCE: 44481-5044-US
; CURRENT APPLICATION NUMBER: US/09/775,803
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: PCT/US99/17594
; PRIOR FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: US 60/109,797
; PRIOR FILING DATE: 1998-08-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-775-803-14
```

```

Query Match          76.2%; Score 1477; DB 10; Length 560;
Best Local Similarity 76.1%; Pred. No. 7.4e-117;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY      1 MLRSALLSAVLALLRAQPPCPKTKCKVVRDAAQCSCGGSVAHIAELGLPTNLTHILFRM 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1 MLRGTLCAVLGLLRAQPPCPPACKCKVFRDAAQCSCGSDVARISALGLPTNLTHILFGM 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      61 DQILRNHSESGMTVLQRLMLSDSHISAIDPGTFNDVLKLTILTRNKISRLPRALDK 120
      :||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



Db 61 GRGVLSQSFSGMTVLQRLMISDSHISAVAPGTFSDLIKLTLSRNKITHLPGALLDK 120  
QY 121 MVLEQLFLDHNAIRDLDONLFPQQLRNLOELGLNQNSFLPANLFSLSRELKLLDLNRN 180  
Db 121 MVLEQLFLDHNAIRGIDQNMFKLVNLOELALNQNLDFLPASLFTNLENKLLDLNRN 180  
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLCALTELRLERNHLRSVAPGAFDR 240  
Db 181 NLTHLPKGLLGAQAKLERLLHSNRLVSLDGLLSNLCALTELQFHRNHIRSIAPGAFDR 240  
QY 241 LGNLSLTLSCNLLLESPLPALFLHVSSVSRLLTLFENPIEELPDVLFEGEMAGRELWLNLT 300  
Db 241 LPNLSLTLSRNHLAFLPSALFLHSNLTLLTLFENPLAELPGVLFEGEMGGLQELWLNRT 300  
QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRVLALHTNALAEIRDDALR 360  
Db 301 QLRTLPAAAFRNLSRLRYLGVTLSPRLSALPQGAFOGLGELQVLALHSNGLTALPDGLLR 360  
QY 361 GLGHLRQVSLRNRRLRALPRTLFRNLSSL 389  
Db 361 GLGKLRQVSLRRNRRLRALPRALFRNLSSL 389

RESULT 7  
US-10-212-499-2  
; Sequence 2, Application US/10212499  
; Publication No. US20030135036A1  
; GENERAL INFORMATION:  
; APPLICANT: Lanza, Francois  
; Phillips, David R.  
; Cazenave, Jean-Pierre  
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan Lewis & Bockius LLP  
; STREET: 1800 M St., NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: US  
; ZIP: 20036  
; COMPUTER READABLE FORM: .  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/212,499  
; FILING DATE: 06-Aug-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/560,814  
; FILING DATE: 2000-04-28  
; APPLICATION NUMBER: US 08/089,455  
; FILING DATE: 1993-07-09  
; APPLICATION NUMBER: US 08/195,006  
; FILING DATE: 1994-02-10  
; APPLICATION NUMBER: US 08/884,571  
; FILING DATE: 1997-06-27  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Reid G. Adler  
; REGISTRATION NUMBER: 30,988  
; REFERENCE/DOCKET NUMBER: 44481-5018-04-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-467-7000  
; TELEFAX: 202-467-7176  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 560 amino  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-212-499-2  
Query Match 76.2%; Score 1477; DB 14; Length 560;

Best Local Similarity 76.1%; Pred. No. 7.4e-117;  
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;  
QY 1 MLRSALLSAVLALLRAQPPCPKCKVVRDAAQCSGGSVAHIAELGLPTNLTHILLPRM 60  
Db 1 MLRGTTLLCAVLGLLRAQPPCPKCKVVRDAAQCSGGSVARISALGLPTNLTHILLPGM 60  
QY 61 DQILRNHSPSGMTVLQRLMISDSHISADPGTFNDLVKLTILRLTRNKISRLPRAILDK 120  
Db 61 GRGVLSQSFSGMTVLQRLMISDSHISAVAPGTFSDLIKLTLSRNKITHLPGALLDK 120  
QY 121 MVLEQLFLDHNAIRDLDONLFPQQLRNLOELGLNQNSFLPANLFSLSRELKLLDLNRN 180  
Db 121 MVLEQLFLDHNAIRGIDQNMFKLVNLOELALNQNLDFLPASLFTNLENKLLDLNRN 180  
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLCALTELRLERNHLRSVAPGAFDR 240  
Db 181 NLTHLPKGLLGAQAKLERLLHSNRLVSLDGLLSNLCALTELQFHRNHIRSIAPGAFDR 240  
QY 241 LGNLSLTLSCNLLLESPLPALFLHVSSVSRLLTLFENPIEELPDVLFEGEMAGRELWLNLT 300  
Db 241 LPNLSLTLSRNHLAFLPSALFLHSNLTLLTLFENPLAELPGVLFEGEMGGLQELWLNRT 300  
QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRVLALHTNALAEIRDDALR 360  
Db 301 QLRTLPAAAFRNLSRLRYLGVTLSPRLSALPQGAFOGLGELQVLALHSNGLTALPDGLLR 360  
QY 361 GLGHLRQVSLRNRRLRALPRTLFRNLSSL 389  
Db 361 GLGKLRQVSLRRNRRLRALPRALFRNLSSL 389

RESULT 8  
US-10-806-018-61  
; Sequence 61, Application US/10806018  
; Publication No. US20040176296A1  
; GENERAL INFORMATION:  
; APPLICANT: Khodadoust, Mehran M.  
; APPLICANT: MacBeth, Kyle J.  
; APPLICANT: Busfield, Samantha J.  
; APPLICANT: McCarthy, Sean A.  
; APPLICANT: Holtzman, Douglas A.  
; APPLICANT: Gu, Wei  
; APPLICANT: White, David  
; APPLICANT: Pan, Yang  
; TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSG, AND  
; TITLE OF INVENTION: STMST PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: MNI-121CP  
; CURRENT APPLICATION NUMBER: US/10/806,018  
; CURRENT FILING DATE: 2004-03-22  
; PRIOR APPLICATION NUMBER: US/09/782,980  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: PCT/US00/02125  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: 09/448,076  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: 09/276,400  
; PRIOR FILING DATE: 1999-03-25  
; PRIOR APPLICATION NUMBER: 60/117,580  
; PRIOR FILING DATE: 1999-01-27  
; PRIOR APPLICATION NUMBER: 09/014,195  
; PRIOR FILING DATE: 1998-01-27  
; PRIOR APPLICATION NUMBER: 09/014,348  
; PRIOR FILING DATE: 1998-01-27  
; PRIOR APPLICATION NUMBER: 09/086,892  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 09/296,208  
; PRIOR FILING DATE: 1999-04-21  
; PRIOR APPLICATION NUMBER: 09/063,950  
; PRIOR FILING DATE: 1998-04-21  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 176









```
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 225
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-080-334-225

Query Match      31.3%; Score 607; DB 15; Length 581;
Best Local Similarity 37.7%; Pred. No. 7.9e-43;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTCCKVVRDAAQCSCGVAHIAELGLPTNLTHILLFRMDQGLRNHSPFGMTVLQRLM 80
Db 25 CPSECTCSRASQVECTGARIVAV-PTPLPWNAMSLOILNTHITELNESPFLNISALIALR 83
QY 81 LSDSHISAIDPGTFNDLVKLTIRLNRKISRPLRAILDKMVLLLEQLFLDHNALRDLDQN 140
Db 84 IEKNELSRITPGAFRNLGSLRYLSLANNKLOVLPGLFQGLDLSLESLLSSNQLLQIOPA 143
QY 141 LFOQLRNQLQELGNQNLQSLFPLPANLFPSSLRRELKLLDLSRNNLTHLPKGLLGAQVKLEKLL 200
Db 144 HFSQCSNLKELQHLGNHLEYIPDGAFDHLVGLTKNLGKNSLTHISPRVFQHLGNLQVLR 203
QY 201 LYSNQLTSVDSGLLSNLGALTELRERNHLRSVAPGAFDRGLNLSLTLSGNLLSPPA 260
Db 204 LYENRLTDIPMGTFDGLVNLQELALQQNQIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263
QY 261 LFLHVSSVSRLLTLFENPLEELPDVLFEGEMAGLRELWNGTHLSTLPAAAFRNLGLQTLG 320
Db 264 IFMQLPQLNRLTLFGNSLKELSGIFGPMPNLRELWLYDNHISLPLDNVFSNLRQLQVLI 323
QY 321 LTRNPRLSALPRGVFQGLRELRLVLAHTNALAELRDDALRGLHRLQVSLRHNRLRALPR 380
Db 324 LSRN-QISFISPGAFNGLTRELRLSLHTNALQDLDDGNVFRMLANLQNISLQNNRLRLQPLG 382
QY 381 TLFRNLSSL 389
Db 383 NIFANVNGL 391

RESULT 14
US-10-258-951-43
; Sequence 43, Application US/10258951
; Publication No. US20040033504A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdock, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; APPLICANT: Lai, Ying-Ta
; APPLICANT: Xie, Qing
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50025
; CURRENT APPLICATION NUMBER: US/10/258,951
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: PCT/US01/13360
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/199,963
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/203,336
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/207,087
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/207,546
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 581
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-258-951-43

Query Match      31.3%; Score 607; DB 15; Length 581;
Best Local Similarity 37.7%; Pred. No. 7.9e-43;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTCCKVVRDAAQCSCGVAHIAELGLPTNLTHILLFRMDQGLRNHSPFGMTVLQRLM 80
Db 25 CPSECTCSRASQVECTGARIVAV-PTPLPWNAMSLOILNTHITELNESPFLNISALIALR 83
QY 81 LSDSHISAIDPGTFNDLVKLTIRLNRKISRPLRAILDKMVLLLEQLFLDHNALRDLDQN 140
Db 84 IEKNELSRITPGAFRNLGSLRYLSLANNKLOVLPGLFQGLDLSLESLLSSNQLLQIOPA 143
QY 141 LFOQLRNQLQELGNQNLQSLFPLPANLFPSSLRRELKLLDLSRNNLTHLPKGLLGAQVKLEKLL 200
Db 144 HFSQCSNLKELQHLGNHLEYIPDGAFDHLVGLTKNLGKNSLTHISPRVFQHLGNLQVLR 203
QY 201 LYSNQLTSVDSGLLSNLGALTELRERNHLRSVAPGAFDRGLNLSLTLSGNLLSPPA 260
Db 204 LYENRLTDIPMGTFDGLVNLQELALQQNQIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263
QY 261 LFLHVSSVSRLLTLFENPLEELPDVLFEGEMAGLRELWNGTHLSTLPAAAFRNLGLQTLG 320
Db 264 IFMQLPQLNRLTLFGNSLKELSGIFGPMPNLRELWLYDNHISLPLDNVFSNLRQLQVLI 323
QY 321 LTRNPRLSALPRGVFQGLRELRLVLAHTNALAELRDDALRGLHRLQVSLRHNRLRALPR 380
Db 324 LSRN-QISFISPGAFNGLTRELRLSLHTNALQDLDDGNVFRMLANLQNISLQNNRLRLQPLG 382
QY 381 TLFRNLSSL 389
Db 383 NIFANVNGL 391

RESULT 15
US-10-080-334-224
; Sequence 224, Application US/10080334
; Publication No. US20040002584A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shinkets, Richard A
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Liu, Xiaohong
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spaderna, Steven K
; APPLICANT: Zerhusen, Bryan D
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-275
; CURRENT APPLICATION NUMBER: US/10/080,334
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,523
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/322,712
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/311,980
```

;; PRIOR FILING DATE: 2001-08-13  
;; PRIOR APPLICATION NUMBER: 60/330,307  
;; PRIOR FILING DATE: 2001-10-18  
;; PRIOR APPLICATION NUMBER: 60/278,796  
;; PRIOR FILING DATE: 2001-03-26  
;; PRIOR APPLICATION NUMBER: 60/281,521  
;; PRIOR FILING DATE: 2001-04-04  
;; PRIOR APPLICATION NUMBER: 60/276,677  
;; PRIOR FILING DATE: 2001-03-16  
;; PRIOR APPLICATION NUMBER: 60/311,595  
;; PRIOR FILING DATE: 2001-08-10  
;; PRIOR APPLICATION NUMBER: 60/270,220  
;; PRIOR FILING DATE: 2001-02-21  
;; PRIOR APPLICATION NUMBER: 60/274,295  
;; PRIOR FILING DATE: 2001-03-08  
;; PRIOR APPLICATION NUMBER: 60/318,526  
;; PRIOR FILING DATE: 2001-09-10  
;; PRIOR APPLICATION NUMBER: 60/286,548  
;; PRIOR FILING DATE: 2001-04-25  
;; PRIOR APPLICATION NUMBER: 60/291,765  
;; PRIOR FILING DATE: 2001-05-17  
;; PRIOR APPLICATION NUMBER: 60/270,797  
;; PRIOR FILING DATE: 2001-02-23  
;; PRIOR APPLICATION NUMBER: 60/276,400  
;; PRIOR FILING DATE: 2001-03-16  
;; PRIOR APPLICATION NUMBER: 60/270,810  
;; PRIOR FILING DATE: 2001-02-23  
;; NUMBER OF SEQ ID NOS: 388  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 224  
;; LENGTH: 787  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-080-334-224

Query Match 31.3%; Score 607; DB 15; Length 787;  
Best Local Similarity 37.7%; Pred. No. 1.2e-42;  
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;  
  
QY 21 CPKCKCVRDAAQCQSGSVAHIAELGLPTNLTHILLFRMDQGLRNHSPSGMTVLQRLM 80  
DB 231 CPSECTCSRASQVECTGARIVAV-PTPLPWAMSLQILNTHITELNESPFNLIALR 289  
  
QY 81 LSDSHISAIDPGTFNDLVKLKTLRLTRNKISRLPRAILDKMVLLLEQLFLDHNAIRDLDQN 140  
DB 290 IEKNELSRITPGAFRNLGSLRYLSLANNKLVLPVIGLFGQLDSLESLLSSNQLLQIPA 349  
  
QY 141 LFOQLRNQLQELGNQNSFLPANFSSRLREIKLLDLSRNNLTHLPKGLLAGQVKEKLL 200  
DB 350 HFSQCSNLKELQLHGNHLEYIPDGAFDHLVGLTKNLGKNSLTHISPRVFOHGLGNLQVLR 409  
  
QY 201 LYSNQLTSVDSGLLSNLGALTTELRLERNHLSRVAPGAFDRGLNLSSTLTSNLLLESPPA 260  
DB 410 LYENRLTDIPMGTFDGLVNLQELALQONQIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 469  
  
QY 261 LFLHVSVSRLLTLFENPLEELPDVLFGEAGRELWLNGLTHSLTPAAAFRNLSGLQTLG 320  
DB 470 IFMQLPQLNRLTLFGNSIKELSLGIFGPMPLRELWLYDNHISISLPDVFNSLRQLQVLI 529  
  
QY 321 LTRNPRLSALPRGVFQGLRELRLVLAHTNALAEALDDALRGHLRQVSLRHNRLRALPR 380  
DB 530 LSRN-QISFISPGAFNGLTRELRLSLHTNALQDLDGNVFRMLANLQNISLQNNRLRLQPG 588  
  
QY 381 TLFRLNSSL 389  
DB 589 NIFANVNGL 597

Search completed: November 29, 2004, 13:18:03  
Job time : 99 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 12:59:28 ; Search time 109 Seconds  
(without alignments)  
2053.401 Million cell updates/sec

Title: US-09-775-803-12\_COPY\_1\_389  
Perfect score: 1938  
Sequence: 1 MLRSALLSAVLALLRAQPP.....LRHNLRLALPRTLFRNLSSL 389

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_02:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1938	100.0	567	2 Q9QZU3	Q9qzu3 mus musculu
2	1923	99.2	567	1 GPV_MOUSE	O08742 mus musculu
3	1749	90.2	567	1 GPV_RAT	O08770 rattus norv
4	1477	76.2	560	1 GPV_HUMAN	P40197 homo sapien
5	782	40.4	637	2 Q6DCV7	Q6dcv7 xenopus lae
6	607	31.3	581	1 LR15_HUMAN	Q8tf66 homo sapien
7	607	31.3	581	2 Q7RTN7	Q7rtn7 homo sapien
8	600	31.0	579	2 Q80X72	Q80x72 mus musculu
9	583	30.1	578	1 LR15_RAT	Q8r5m3 rattus norv
10	541.5	27.9	391	2 Q9D3K0	Q9d3k0 m mus muscu
11	503.5	26.0	562	2 Q8R113	Q8r1l3 mus musculu
12	503.5	26.0	570	2 Q9DEB9	Q9dbb9 mus musculu
13	445	23.0	557	2 Q8N5V4	Q8n5v4 homo sapien
14	445	23.0	560	2 Q86SU4	Q86su4 homo sapien
15	419.5	21.6	605	1 ALS_HUMAN	P35858 homo sapien
16	416.5	21.5	536	1 CBP8_HUMAN	P22792 homo sapien
17	416.5	21.5	605	2 Q8TAY0	Q8tay0 homo sapien
18	413.5	21.3	605	1 ALS_PAPHA	O02833 papio hamad
19	397	20.5	3127	2 Q7PTD4	Q7ptd4 anopheles g
20	384	19.8	1000	2 Q7QF88	Q7qf88 anopheles g
21	382.5	19.7	603	1 ALS_RAT	P35859 rattus norv
22	382	19.7	603	1 ALS_MOUSE	P70389 mus musculu
23	382	19.7	603	2 AAF69482	Aaf69482 mus muscu
24	382	19.7	687	2 Q9JIL0	Q9jil0 mus musculu
25	381.5	19.7	603	2 Q70211	Q70211 rattus norv
26	379.5	19.6	1531	1 SLT1_RAT	O88279 rattus norv
27	379.5	19.6	1534	1 SLT1_HUMAN	Q75093 homo sapien
28	379	19.6	417	2 Q6E4J7	Q6e4j7 petromyzon
29	375.5	19.4	1531	1 SLT1_MOUSE	Q80tr4 mus musculu
30	372.5	19.2	1058	2 Q76FN7	Q76fn7 tachypleus
31	372.5	19.2	1058	2 BAD12073	Bad12073 tachypleu

32	370	19.1	890	2 Q7Q941	Q7q941 anopheles g
33	369.5	19.1	766	1 SLT2_RAT	Q9wvc1 rattus norv
34	367.5	19.0	907	1 LGR5_MOUSE	Q9zlp4 mus musculu
35	367.5	19.0	1521	1 SLT2_MOUSE	Q9rlb9 mus musculu
36	366.5	18.9	1392	2 Q9VAD1	Q9vad1 drosophila
37	366	18.9	1523	1 SLT3_RAT	O88280 rattus norv
38	365	18.8	451	2 Q7QIS1	Q7qis1 anopheles g
39	364.5	18.8	615	2 Q9VZ84	Q9vz84 drosophila
40	364.5	18.8	1529	2 Q7ZX12	Q7zxi2 xenopus lae
41	362	18.7	1523	1 SLT3_MOUSE	Q9wvb4 mus musculu
42	361.5	18.7	1529	1 SLT2_HUMAN	O94813 homo sapien
43	361.5	18.7	1530	2 Q90WZ3	Q90wz3 xenopus lae
44	361	18.6	1173	2 Q7QHH1	Q7qhh1 anopheles g
45	360	18.6	737	2 Q965M3	Q965m3 caenorhabdi

ALIGNMENTS

RESULT 1  
Q9QZU3  
ID Q9QZU3 PRELIMINARY; PRT; 567 AA.  
AC Q9QZU3;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)  
DE Platelet glycoprotein V (Fragment) (Mus musculus 13 days embryo male  
testis cDNA, RIKEN full-length enriched library, clone:6030400G03  
DE product:glycoprotein 5 (platelet), full insert sequence)  
DE (Fragment).  
GN Name=Gp5;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RX MEDLINE=20027548; PubMed=10557321;  
RA Ramakrishnan V., Reeves P.S., DeGuzman F., Deshpande U.,  
Ministri-Madrid K., DuBridge R.B., Phillips D.R.;  
RT "Increased thrombin responsiveness in platelets from mice lacking  
glycoprotein V."; Proc. Natl. Acad. Sci. U.S.A. 96:13336-13341(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=99279253; PubMed=10349636;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning."; Meth. Enzymol. 303:19-44(1999).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=21085660; PubMed=11217851;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RA The FANTOM Consortium,  
RT the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs."; Nature 420:563-573(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=20499374; PubMed=11042159;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [6]

RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=20530913; PubMed=11076861;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [7]

RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF163101; AAF08787.1; -;  
DR EMBL; AK031272; BAC27331.1; -;  
DR HSSP; Q9BZR6; 10ZN.  
DR MGD; MGI:1096363; Gp5.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR003591; LRR\_Typ.  
DR Pfam; PF00560; LRR; 14.  
DR Pfam; PF01463; LRRCT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00369; LRR\_TYP; 10.  
FT NON TER 567  
SQ SEQUENCE 567 AA; 63381 MW; E6AF470855ACB6C9 CRC64;

Query Match 100.0%; Score 1938; DB 2; Length 567;  
Best Local Similarity 100.0%; Pred. No. 2.3e-122;  
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRSALLSAVALLRAQPPCPKTKCVVRDAAQCSCGSGVAHAEGLPTNLTHILLFRM 60  
DB 1 MLRSALLSAVALLRAQPPCPKTKCVVRDAAQCSCGSGVAHAEGLPTNLTHILLFRM 60  
QY 61 DQGLRNHFSFGMTVLQRLMLSDSHISAIDPGTFENDLVKLTLETRNKISRLPRAILDK 120  
DB 61 DQGLRNHFSFGMTVLQRLMLSDSHISAIDPGTFENDLVKLTLETRNKISRLPRAILDK 120  
QY 121 MVLLQLFLDHNALRDLDQNLFOQLRNQLQELGLNQNLQSLFPLANFSSRLRELKLLDLSRN 180  
DB 121 MVLLQLFLDHNALRDLDQNLFOQLRNQLQELGLNQNLQSLFPLANFSSRLRELKLLDLSRN 180  
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTLELERNHLRSVAPGFDR 240  
DB 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTLELERNHLRSVAPGFDR 240  
QY 241 LGNLSLTLTSGNLLSLPALFLHVSSVSRITLFFENPLEELPDVLFEGMAGLRELWNGT 300  
DB 241 LGNLSLTLTSGNLLSLPALFLHVSSVSRITLFFENPLEELPDVLFEGMAGLRELWNGT 300  
QY 301 HLSTLPAAAFRLNSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAELRDDALR 360  
DB 301 HLSTLPAAAFRLNSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAELRDDALR 360

QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389  
DB 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389  
RESULT 2  
GPV\_MOUSE  
ID GPV\_MOUSE STANDARD; PRT; 567 AA.  
AC O08742;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Platelet glycoprotein V precursor (GPV) (CD42D).  
GN Name=Gp5;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Liver;  
RX MEDLINE=97275136; PubMed=9129030;  
RA Ravanat C., Morales M., Azorsa D.O., Moog S., Schuhler S., Grunert P.,  
RA Loew D., van Dorsselaer A., Cazenave J.-P., Lanza F.;  
RT "Gene cloning of rat and mouse platelet glycoprotein V: identification  
of megakaryocyte-specific promoters and demonstration of functional  
thrombin cleavage.";  
RL Blood 89:3253-3262(1997).  
CC -!- FUNCTION: The GPIb-V-IX complex functions as the von Willebrand  
factor receptor and mediates von Willebrand factor-dependent  
platelet adhesion to blood vessels. The adhesion of platelets to  
injured vascular surfaces in the arterial circulation is a  
critical initiating event in hemostasis (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- SIMILARITY: Contains 14 leucine-rich (LRR) repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; Z69595; CAA93441.1; -;  
DR HSSP; Q9BZR6; 10ZN.  
DR MGD; MGI:1096363; Gp5.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR003591; LRR\_Typ.  
DR Pfam; PF00560; LRR; 13.  
DR Pfam; PF01463; LRRCT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00369; LRR\_TYP; 10.  
KW Blood coagulation; Cell adhesion; Glycoprotein; Leucine-rich repeat;  
KW Platelet; Repeat; Signal; Transmembrane.  
FT SIGNAL 1 16 Potential.  
FT CHAIN 17 567 Platelet glycoprotein V.  
FT DOMAIN 17 522 Extracellular (Potential).  
FT TRANSMEM 523 543 Potential.  
FT DOMAIN 544 567 Cytoplasmic (Potential).  
FT REPEAT 73 96 LRR 1.  
FT REPEAT 97 120 LRR 2.  
FT REPEAT 122 144 LRR 3.  
FT REPEAT 145 168 LRR 4.  
FT REPEAT 170 192 LRR 5.  
FT REPEAT 194 216 LRR 6.  
FT REPEAT 217 240 LRR 7.  
FT REPEAT 241 264 LRR 8.  
FT REPEAT 266 288 LRR 9.  
FT REPEAT 289 312 LRR 10.





QY 241 LGNLSLTSGNLLSPPALFLHVSSVRLTLFENPLELPDVLFGEMAGLRELWNGT 300  
Db 241 LGNLSLTSGNLLSPPALFLHVSWLRLTLFENPLELPDVLFGEMAGLRELWNGT 300  
QY 301 HUSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAELEDDALR 360  
Db 301 HURTLPAAFRNLSGLQTLGLTRNPLLSALPPGMFHLGTLRLVLAHTNALAELEDPEDALR 360  
QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389  
Db 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389

## RESULT 4

GPV\_HUMAN STANDARD; PRT; 560 AA.  
AC P40197;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Platelet glycoprotein V precursor (GPV) (CD42D).  
GN Name=GP5;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=93391348; PubMed=7690959;  
RA Hickey M.J., Hagen F.S., Yagi M., Roth G.J.;  
RT "Human platelet glycoprotein V: characterization of the polypeptide  
and the related Ib-V-IX receptor system of adhesive, leucine-rich  
glycoproteins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:8327-8331(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Platelet;  
RX MEDLINE=94012616; PubMed=8407908;  
RA Lanza F., Morales M., de la Salle C., Cazenave J.-P., Clemetson K.J.,  
RA Shimomura T., Phillips D.R.;  
RT "Cloning and characterization of the gene encoding the human platelet  
glycoprotein V. A member of the leucine-rich glycoprotein family  
cleaved during thrombin-induced platelet activation.";  
RL J. Biol. Chem. 268:20801-20807(1993).  
RN [3]  
RP PARTIAL SEQUENCE.  
RC TISSUE=Platelet;  
RX MEDLINE=90275263; PubMed=2350580;  
RA Shimomura T., Fujimura K., Maehama S., Takemoto M., Oda K.,  
RA Fujimoto T., Oyama R., Suzuki M., Ichihara-Tanaka K., Titani K.,  
RA Kuramoto A.;  
RT "Rapid purification and characterization of human platelet  
glycoprotein V: the amino acid sequence contains leucine-rich  
repetitive modules as in glycoprotein Ib.";  
RL Blood 75:2349-2356(1990).  
RN [4]  
RP PARTIAL SEQUENCE.  
RC TISSUE=Platelet;  
RX MEDLINE=90321220; PubMed=2372284;  
RA Roth G.J., Church T.A., McMullen B.A., Williams S.A.;  
RT "Human platelet glycoprotein V: a surface leucine-rich glycoprotein  
related to adhesion.";  
RL Biochem. Biophys. Res. Commun. 170:153-161(1990).  
CC -!- FUNCTION: The GPIb-V-IX complex functions as the von Willebrand  
factor receptor and mediates von Willebrand factor-dependent  
platelet adhesion to blood vessels. The adhesion of platelets to  
injured vascular surfaces in the arterial circulation is a  
critical initiating event in hemostasis.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Platelets and megakaryocytes.  
CC -!- PTM: The N-terminus is blocked.

CC -!- SIMILARITY: Contains 14 leucine-rich (LRR) repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; L11238; AAA03069.1; -;  
DR EMBL; Z23091; CAA80637.1; -;  
DR PIR; A48030; A60164.  
DR HSSP; P07359; 1M0Z.  
DR Genew; HGNC:4443; GP5.  
DR MIM; 173511; -;  
DR GO; GO:0005887; C: integral to plasma membrane; TAS.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR003591; LRR\_typ.  
DR Pfam; PF00560; LRR; 14.  
DR Pfam; PF01463; LRRCT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
KW Blood coagulation; Cell adhesion; Direct protein sequencing;  
KW Glycoprotein; Leucine-rich repeat; Platelet; Repeat; Signal;  
KW Transmembrane.  
FT SIGNAL 1 16 Potential.  
FT CHAIN 17 560 Platelet glycoprotein V.  
FT DOMAIN 17 523 Extracellular (Potential).  
FT TRANSMEM 524 544 Potential.  
FT DOMAIN 545 560 Cytoplasmic (Potential).  
FT REPEAT 73 96 LRR 1.  
FT REPEAT 97 120 LRR 2.  
FT REPEAT 122 144 LRR 3.  
FT REPEAT 145 168 LRR 4.  
FT REPEAT 169 192 LRR 5.  
FT REPEAT 194 216 LRR 6.  
FT REPEAT 217 240 LRR 7.  
FT REPEAT 241 264 LRR 8.  
FT REPEAT 266 288 LRR 9.  
FT REPEAT 289 312 LRR 10.  
FT REPEAT 314 337 LRR 11.  
FT REPEAT 338 361 LRR 12.  
FT REPEAT 362 385 LRR 13.  
FT REPEAT 386 409 LRR 14.  
FT CARBOHYD 51 51 N-linked (GlcNAc...).  
FT CARBOHYD 181 181 N-linked (GlcNAc...).  
FT CARBOHYD 243 243 N-linked (GlcNAc...).  
FT CARBOHYD 267 267 N-linked (GlcNAc...).  
FT CARBOHYD 298 298 N-linked (GlcNAc...).  
FT CARBOHYD 312 312 N-linked (GlcNAc...).  
FT CARBOHYD 385 385 N-linked (GlcNAc...).  
FT CARBOHYD 499 499 N-linked (GlcNAc...).  
FT CONFLICT 73 74 MT -> TK (in Ref. 2).  
FT CONFLICT 109 109 K -> T (in Ref. 2).  
FT CONFLICT 130 130 D -> W (in Ref. 3).  
FT CONFLICT 136 138 GID -> PGG (in Ref. 3).  
FT CONFLICT 209 209 L -> I (in Ref. 2).  
FT CONFLICT 267 267 N -> H (in Ref. 3).  
FT CONFLICT 327 327 L -> I (in Ref. 2).  
FT CONFLICT 478 478 P -> G (in Ref. 2).  
FT CONFLICT 509 509 P -> D (in Ref. 2).  
SQ SEQUENCE 560 AA; 60959 MW; 'BICDB04AF8AF7115 CRC64;

Query Match 76.2%; Score 1477; DB 1; Length 560;  
Best Local Similarity 76.1%; Pred. No. 2.7e-91;  
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 1 MLRSALLSAVLALLRAQPPPCPKCKVVRDAAQCSGVAHIAELGLPTNLTHILLFRM 60  
Db 1 MLRGTLCAVLGLLRAQPPPCPKCKVVRDAAQCSGVAHIAELGLPTNLTHILLFRM 60  
QY 61 DQGILRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKISRLPAIDK 120

Db 61 GRCVLSQSFSGMTVLQRLMISDSHISAVAPGTFSDLIKLTLSRNKITHLPGLLDK 120  
QY 121 MVLEQLFELDHNAIRDLDONLFQQLRNQLQELGLNQNLQSLPANLFSLSRELKLLDLRN 180  
Db 121 MVLEQLFELDHNAIRDLDONLFQQLRNQLQELGLNQNLQSLPANLFSLSRELKLLDLRN 180  
QY 181 NLTHLPKGLGCAQVKLEKLLLYNQNLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDR 240  
Db 181 NLTHLPKGLGCAQVKLEKLLLYNQNLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDR 240  
QY 241 LGNLSLTLSCNLLSLPPLFELHVSVSRLTLFENPLELDPVLFEGMAGLRELWLNGT 300  
Db 241 LPNLSLTLSCNLLSLPPLFELHVSVSRLTLFENPLELDPVLFEGMAGLRELWLNGT 300  
QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAELEDDALR 360  
Db 301 QLRTPAAAFRNLSRLRYLGVTLSPRLSALPQGAFOGLGELQVLALHSNGLTALPDGLLR 360  
QY 361 GLGHLRQVSLRNRRLRALPRTLFRNLSSL 389  
Db 361 GLGHLRQVSLRNRRLRALPRTLFRNLSSL 389

RESULT 5  
Q6DCV7 PRELIMINARY; PRT; 637 AA.  
ID Q6DCV7  
AC Q6DCV7  
DT 01-OCT-2004 (Tremblrel. 28, Created)  
DT 01-OCT-2004 (Tremblrel. 28, Last sequence update)  
DT 01-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE Hypothetical protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX MEDLINE=22341132; PubMed=12454917;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.,  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative."  
RL Dev. Dyn. 225:384-391 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.C., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;

RA Klein S., Strausberg R.,  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC077882; AAH77882.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 637 AA; 72764 MW; 2D174E8895789F23 CRC64;  
Query Match 40.4%; Score 782; DB 2; Length 637;  
Best Local Similarity 45.3%; Pred. No. 2.3e-44;  
Matches 167; Conservative 58; Mismatches 144; Indels 0; Gaps 0;  
QY 21 CPKTCVCVVRDAAQCSSGVAHIAELGLPTNLTHILLFRMDQILRNHSFGMTVLQRLM 80  
Db 18 CPTLCTCRKDAVFCQGPSIKDIGSLLPSNFTYIHIINTLATEITDKSFGNMPITLRLR 77  
QY 81 LSDSHISAIDPGTFNDLVKLTILTRNKRISRLPRAILLDRXVLLLEQLFLDHNAIRDLDQN 140  
Db 78 LEDSRLTFTTRDAFKSLPQLKSLKLTNNKLETLPAGVFDLSLFYLEQLFIGNHLSLHPN 137  
QY 141 LFQQLRNQLGELNQNLQSLPANLFSLSRELKLLDLRSNNLTHLPKGLLGAQVKLEKLL 200  
Db 138 LFCCLQHLKELILNRNQNLTSPLNELRLNLTITLNLNRNKISHLPVSIFSSLTCLKKLH 197  
QY 201 LYSNQLTSVDSGLLSNLGALTTELRLERNHLRSVAPGAFDRIGLNLSSLTLSCNLLSLPPA 260  
Db 198 LYENQLTITSSAFNNLGELELALYSNSIQSIAPDAFHLPKLLNLSXNKLHFLPYG 257  
QY 261 LFLHVSVSRLTLFENPLELDPVLFEGMAGLRELWLNGTHLSTLPAAAFRNLSGLQTLG 320  
Db 258 LFLHLPQLSVLTLYDNPLKELPVDVIFGKMENLTSLWLYDTHLATIPNFVFCNLTNQLLV 317  
QY 321 LTRNPRLSALPRGVFQGLRELRLVLAHTNALAELEDDALRGLRQVSLRNRRLRALPR 380  
Db 318 LTRNPQLESPLADAFSGLSNLELSLHNSNLSIDQLFQNLQLEKLSLYSNLKVLS 377  
QY 381 TLFRNLSSL 389  
Db 378 NMFYNLNL 386

RESULT 6  
LR15\_HUMAN  
ID LR15\_HUMAN STANDARD; PRT; 581 AA.  
AC Q8TF66;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Leucine-rich repeat-containing protein 15 precursor (hLib).  
GN Name=LRRCL15; Synonyms=LIB;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=21645900; PubMed=11785964; DOI=10.1006/bbrc.2001.6272;  
RA Satoh K., Hata M., Yokota H.;  
RT "A novel member of the leucine-rich repeat superfamily induced in rat  
RT astrocytes by beta-amyloid."  
RL Biochem. Biophys. Res. Commun. 290:756-762(2002).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
CC -!- TISSUE SPECIFICITY: Brain and placenta.  
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
DR EMBL; AB071037; BAB84587.1; -.



DR HSP; Q9BZR6; 10ZN.  
DR Genew; HGNC:20818; LRRCL5.  
DR InterPro; IPR011061; Antihemostatic.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003591; LRR\_typ.  
DR Pfam; PF00560; LRR; 14.  
DR PRINTS; PRO0019; LEURICHRPT.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_TYP; 10.  
KW Leucine-rich repeat; Repeat; Signal; Transmembrane.  
FT SIGNAL 1 21  
FT CHAIN 22 581  
FT  
FT DOMAIN 22 538  
FT TRANSMEM 539 559  
FT DOMAIN 560 581  
FT REPEAT 51 75  
FT REPEAT 76 99  
FT REPEAT 100 123  
FT REPEAT 125 147  
FT REPEAT 148 171  
FT REPEAT 173 195  
FT REPEAT 196 219  
FT REPEAT 221 243  
FT REPEAT 244 267  
FT REPEAT 269 291  
FT REPEAT 292 315  
FT REPEAT 317 339  
FT REPEAT 340 363  
FT REPEAT 364 387  
FT REPEAT 389 411  
FT CARBOHYD 75  
FT CARBOHYD 369  
FT SEQUENCE 581 AA; 64396 MW; 1F381485BD2CCB54 CRC64;  
SQ

Query Match 31.3%; Score 607; DB 1; Length 581;  
Best Local Similarity 37.7%; Pred. No. 1.3e-32;  
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKCKCVRDAAQCSCGSAHIAELGLPTNLTHILLFRMDQGLRNHSPFGMTVLQRLM 80  
DB 25 CPSECTCSRASQVECTGARIVAV-PTPLPWAMSLQILNTHITELNESPFLNISALIALR 83  
QY 81 LSDSHISAIDPGTFNDLVKLTLRTRNKISRLPRAILDKNVLLLEQLFLDHNALDLDQN 140  
DB 84 IEKNELSRITPGAFRLNLSRLYLSLANNKLVLPGLFGQLDSLESLLSSNQLLIQIPA 143  
QY 141 LFOQLRNLOELGLNQSLFPLANLFSLSRELKLLDLRSNNLTHLPKGLLGQAQVKLEKLL 200  
DB 144 HFSQCSNLKELQHGHNLEYIPDGAFDHLVGLTKLNKNSLTHISPRVFOHGLNQLVLR 203  
QY 201 LYSNQLTSVDSGLLSNIGALTELRLERNHLSVAPGAFDRLGNLSSLTLSGNLLESPLPA 260  
DB 204 LYENRLTDIPMGTFDGLVNLQELALQONQIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263  
QY 261 LFLHVSSVSRITLFPENPLEELPDVLFEGMAGRELWNGTHLSTLPAAAFRLNLSGLQTIG 320  
DB 264 IFMQLPQLNRLTLFGNSLKELSLGIFGPMPLNRELWLYDNHISLPLDNLVFNLSQLPPL 323  
QY 321 LTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALRGHLRQVSLRHNELRALPR 380  
DB 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDLGDNVFRMLANLQNISLQNNELRLQLPG 382  
QY 381 TLFRNLSSL 389  
DB 383 NIFANVNGL 391

RESULT 7  
Q7RTN7

ID Q7RTN7 PRELIMINARY; PRT; 581 AA.  
AC Q7RTN7;  
DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE LRRCL5.  
GN Name=LRRCL5;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22833479; PubMed=12923058;  
RA Reynolds P.A., Smolen G.A., Palmer R.E., Sgroi D., Yajnik V.,  
RA Gerald W.L., Haber D.A.;  
RT "Identification of a DNA-binding site and transcriptional target for  
RT the EWS-WT1(+KTS) oncoprotein.";  
RL Genes Dev. 17:2094-2107(2003).  
CC -!- MISCELLANEOUS: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.  
DR EMBL; BK001325; DAA01740.1; -.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR003591; LRR\_typ.  
DR Pfam; PF00560; LRR; 14.  
DR PRINTS; PRO0019; LEURICHRPT.  
SQ SEQUENCE 581 AA; 64396 MW; 1F381485BD2CCB54 CRC64;

Query Match 31.3%; Score 607; DB 2; Length 581;  
Best Local Similarity 37.7%; Pred. No. 1.3e-32;  
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKCKCVRDAAQCSCGSAHIAELGLPTNLTHILLFRMDQGLRNHSPFGMTVLQRLM 80  
DB 25 CPSECTCSRASQVECTGARIVAV-PTPLPWAMSLQILNTHITELNESPFLNISALIALR 83  
QY 81 LSDSHISAIDPGTFNDLVKLTLRTRNKISRLPRAILDKNVLLLEQLFLDHNALDLDQN 140  
DB 84 IEKNELSRITPGAFRLNLSRLYLSLANNKLVLPGLFGQLDSLESLLSSNQLLIQIPA 143  
QY 141 LFOQLRNLOELGLNQSLFPLANLFSLSRELKLLDLRSNNLTHLPKGLLGQAQVKLEKLL 200  
DB 144 HFSQCSNLKELQHGHNLEYIPDGAFDHLVGLTKLNKNSLTHISPRVFOHGLNQLVLR 203  
QY 201 LYSNQLTSVDSGLLSNIGALTELRLERNHLSVAPGAFDRLGNLSSLTLSGNLLESPLPA 260  
DB 204 LYENRLTDIPMGTFDGLVNLQELALQONQIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263  
QY 261 LFLHVSSVSRITLFPENPLEELPDVLFEGMAGRELWNGTHLSTLPAAAFRLNLSGLQTIG 320  
DB 264 IFMQLPQLNRLTLFGNSLKELSLGIFGPMPLNRELWLYDNHISLPLDNLVFNLSQLPPL 323  
QY 321 LTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALRGHLRQVSLRHNELRALPR 380  
DB 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDLGDNVFRMLANLQNISLQNNELRLQLPG 382  
QY 381 TLFRNLSSL 389  
DB 383 NIFANVNGL 391

RESULT 8  
Q80X72

ID Q80X72 PRELIMINARY; PRT; 579 AA.  
AC Q80X72;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Lrrc15 protein.  
GN Name=Lrrc15;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.









Genome Res. 10:1617-1630(2000).  
[5]  
SEQUENCE FROM N.A.  
STRAIN=C57BL/6J; TISSUE=Liver;  
MEDLINE=20530913; PubMed=11076861;  
RA Shibata K., Itoh M., Aizawa K., Kitsuai T., Tashiro H., Itoh M.,  
RA Konno H., Akiyama J., Nishi K., Kitsumai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
[6]  
SEQUENCE FROM N.A.  
STRAIN=C57BL/6J; TISSUE=Liver;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.;  
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AK005049; BAB23775.1; -.  
DR HSSP; P07359; 1M0Z.  
DR MGD; MGI:1919006; 1300018K11rik.  
DR GO; GO:0004180; F:carboxypeptidase activity; IEA.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003591; LRR\_typ.  
DR Pfam; PF00560; LRR; 12.  
DR Pfam; PF01462; LRRNT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_typ; 8.  
KW Carboxypeptidase.  
SQ SEQUENCE 570 AA; 62889 MW; 634CEE2B5F467C7F CRC64;  
  
Query Match 26.0%; Score 503.5; DB 2; Length 570;  
Best Local Similarity 34.6%; Pred. No. 1.2e-25;  
Matches 129; Conservative 69; Mismatches 172; Indels 3; Gaps 3;  
  
QY 11 LALLRAQFPCKTKCVVRDAAQCSCGGSVAHIAELGLPTNLTHILLFRMDQGLNHSF 70  
Db LLLARLTQPCVGCDFGREVF-CSDEQLADIPP-DIPPHITDIVFVETAFITVTRAF 92  
QY 71 SGMVTLQRLMLSDSHISAIDPGTFNDLVKLKTLRLTRNKISRLPRAILDKNVLLLEQLFD 130  
Db SGPNTLTKVFLNTQVRHLEPDAGFGLPRQDLEITGSPVSNLSAHIFSNLSLEKLTLD 152  
QY 131 HNALRDLDQNLFQQLRNQLGELGNQNLFLPANLFSLSRELKLLDLSRNNLTHLPKGLL 190  
Db FRLAGLPEDLFCHMDILESLOLQGNQLTLPGRLFQSLRDLRTLNLAQNLLTQPKGAP 212  
QY 191 GAQVKLEKLLYSNQLTSVDSGLLSNIGALTELRLERNHLRSVAPGAFDRGLNLSLTL 250  
Db QSLTGLQMLKLSNNMLARLPEGALGSLSSQLFELFDGNAITELSPHLSQLFSLEMLWQ 272  
QY 251 GNLESPPALFLHVSSVSRLTLFENPLEELPDVLFGEAGRELWNGTHLSLTPAAAF 310  
Db HNAICHLPVSLFSSSLHNLTLFLSKDNALRTLPEGLFAHNOGLHLHLSLNSYNQLETFEGAF 332  
QY 311 RNLSGLQTLGLTRNPRLSALPRGVFGQLRELRLVLAHTNALAEALRDLALRGLHRLQVSL 370

Db 333 TNLRLVSLTSLSH-AITDLPEHVRNLEQLVKLSLSDSNLTALHPALFHNLSRLQLNL 391  
QY 371 RHNELRALPRTL 383  
Db 392 SRNQLTTLPGGIF 404  
  
RESULT 13  
Q8NSV4  
ID Q8NSV4 PRELIMINARY; PRT; 557 AA.  
AC Q8NSV4;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE Similar to RIKEN CDNA 1300018K11 gene (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon;  
RA Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC031569; AAH31569.1; -.  
DR HSSP; Q9BZR6; 1P8T.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003591; LRR\_typ.  
DR Pfam; PF00560; LRR; 13.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01462; LRRNT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_typ; 10.  
FT NON TER 1  
SQ SEQUENCE 557 AA; 61782 MW; A65BFC340CC76563 CRC64;  
  
Query Match 23.0%; Score 445; DB 2; Length 557;  
Best Local Similarity 33.1%; Pred. No. 1e-21;  
Matches 125; Conservative 54; Mismatches 173; Indels 26; Gaps 4;  
  
QY 1 MLRSALLSAVLALRAQPF-PCPKTKCVVRDAAQCSCGGSVAHI----- 43  
Db MLPGAWLLWTSLLARPAQPCPMGDCFCVQEVF-CSDEELATVPLDIPPYTKNIIFVET 71  
QY 44 -----AELGLPTNLTHILLFRMDQGLRNHSFSGMTVLQRLMLSDSHISAIDPGTFND 96  
Db SFTTLETRAFGSPNLTWKVFLNTQLCCQPRPDAGFGLPRLEDEVTGSSFLNLSNIFSN 131  
QY 97 LVKLKTLRLTRNKISRLPRAILDKNVLLLEQLFDHNAIRDLPQNLFQQLRNLOELGNQN 156  
Db LTSGLKLTNLFNNMLEALPEGLFQHLAALLESLLHQQNLQALPRRLFPQLTHLKTNLNAQN 191  
QY 157 QLSFLPANLFSLSRELKLLDLSRNNLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSN 216  
Db LLAQLPEELFHPLTSLQTLKLSNNALSGLPQGVFGKLGSLQELFSDSNISSELPQVFSQ 251  
QY 217 LGALTELRLERNHLRSVAPGAFDRGLNLSLTLSGNLLSLELPPALFLHVSSVSRLTFEN 276  
Db LFCLERLWLQRNAITHLPLSIFASLGNLTFLSLQWNMLRVLPAGLFAHTPCLVGLSLTHN 311  
QY 277 PLEELPDVLFGEAGRELWNGTHLSLTPAAAFRLNSGLQTLGLTRNPRLSALPRGVFG 336  
Db QLETVAGETFAHLSNLSRLSLSYNAITHLDPAGIFRDLBELVKLYLGSN-NLTALHPLFQ 370  
QY 337 GRELRLVLAHTNALAE 354  
Db NLSKLELLSLSKNQLTTL 388



RESULT 14

Q86SU4 PRELIMINARY; PRT; 560 AA.  
AC Q86SU4;  
DT 01-JUN-2003 (TReMBLrel. 24, Created)  
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
DE Similar to RIKEN cDNA 1300018K11 gene (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon;  
RA Strausberg R.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC042334; AAH42334.1; -  
DR HSSP; Q9BZR6; LOZN.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003591; LRR\_typ.  
DR Pfam; PF00560; LRR; 13.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01462; LRRNT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_TYP; 10.  
FT NON TER 1  
SQ SEQUENCE 560 AA; 62196 MW; CABFEBCL1E69B0934 CRC64;

RESULT 15

	ID	ALS HUMAN AC	STANDARD;	PRT;	605 AA.
DT	01-JUN-1994	(Rel. 29, Created)			
DT	01-JUN-1994	(Rel. 29, Last sequence update)			
DT	05-JUL-2004	(Rel. 44, Last annotation update)			

FT	REPEAT	74	96	LRR 2.
FT	REPEAT	98	120	LRR 3.
FT	REPEAT	121	144	LRR 4.
FT	REPEAT	145	168	LRR 5.
FT	REPEAT	170	192	LRR 6.
FT	REPEAT	193	216	LRR 7.
FT	REPEAT	217	240	LRR 8.
FT	REPEAT	242	264	LRR 9.
FT	REPEAT	265	288	LRR 10.
FT	REPEAT	289	312	LRR 11.
FT	REPEAT	313	336	LRR 12.
FT	REPEAT	337	360	LRR 13.
FT	REPEAT	361	384	LRR 14.
FT	REPEAT	386	408	LRR 15.
FT	REPEAT	409	432	LRR 16.
FT	REPEAT	433	456	LRR 17.
FT	REPEAT	458	480	LRR 18.
FT	REPEAT	482	504	LRR 19.
FT	REPEAT	505	530	LRR 20.
FT	CARBOHYD	64	64	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	85	85	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	96	96	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	368	368	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	515	515	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	580	580	N-linked (GlcNAc. . .) (Potential).
SQ	SEQUENCE	605 AA; 66034 MW; F6562A23CBE918F6 CRC64;		
Query Match 21.6%; Score 419.5; DB 1; Length 605;				
Best Local Similarity 33.0%; Pred. No. 5.9e-20;				
Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;				
QY	38 GSAH--IAELGLPTNLTHILLFRMDQGI LRHSFGMTVLQRLMLSDSHISAIDPGTF 94			
Db	140 GTPAHTPALASGLSNN-----RLSRLEDGL-----FEGLSLWDNLGWNLSLAVLPDAAF 190			
QY	95 NDLVKLKTLRTRNKISRLPRAILDKMVLLEQLFLDHNAIRDLDQNLFOQLRNLQELGLN 154			
Db	191 RGLGSLRELVLGNRLAYLQPALFSGLAELRELDLSRNALRAIKANVFVQLPRLQKLYLD 250			
QY	155 QNQLSFLPANFSSRELKLLDLSNNLTHLPK----GLCAQVKLEKLLLYSNQLTSVD 210			
Db	251 RNLIAAVAPGAFGLKALRWLDLSHNRVAGLLEDTFPGLLGLRV----LRLSHNAIASLR 306			
QY	211 SGLLSNLGALTELRERLERNHLSRVAPGAPDRGLNLSLTLSCNLLLESPLPPALFHVSSVSR 270			
Db	307 PRTEKDLHFLEELQLGHNRI RQLAERSFEGLGQLEVLTLTDHNLQAEVKAGAFGLTNAVAV 366			
QY	271 LTLFENPLEELPDVLFGE MAGRLWLNGLTHLSTLPAAAFRNLSGLQTLGLTRN----- 324			
Db	367 MNLSGNCRLNLPQVFRGLGKLHSLHLEGSCLGRIRPHFTTGLSGLRLFLKDNGLVGIE 426			
QY	325 -----PRLSALPRGVFQGLRELRLVLAHTNALAELRDDALRGLGHLRQ 367			
Db	427 EQSLWGLAELELDLTSNQLTHPLRFLQGLKLEYLLLSRNRLAELPADALGPLQRAFW 486			
QY	368 VSLRNRRLRALPRTLFRNLSSL 389			
Db	487 LDVSHNRLEALPNSLLAPLGR 508			

Search completed: November 29, 2004, 13:10:59  
Job time : 113 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 13:05:23 ; Search time 27 Seconds  
(without alignments)  
955.470 Million cell updates/sec

Title: US-09-775-803-12\_COPY\_1\_389  
Perfect score: 1938  
Sequence: 1 MLRSALLSAVLALRAQPPF.....LRHRLRALPRTLFRNLSSL 389

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1477	76.2	560	3	US-08-592-500-2
2	1477	76.2	560	3	US-08-195-006-2
3	1477	76.2	560	3	US-09-063-950-4
4	1477	76.2	560	5	PCT-US94-07644A-2
5	419.5	21.6	605	1	US-08-190-802A-49
6	419.5	21.6	605	3	US-08-477-346-49
7	419.5	21.6	605	3	US-08-473-089-49
8	419.5	21.6	605	4	US-08-487-072A-49
9	419.5	21.6	605	4	US-09-538-092-1087
10	416.5	21.5	536	4	US-09-538-092-992
11	413.5	21.3	605	3	US-09-063-950-5
12	382.5	19.7	603	1	US-08-190-802A-50
13	382.5	19.7	603	3	US-08-477-346-50
14	382.5	19.7	603	3	US-08-473-089-50
15	382.5	19.7	603	4	US-08-487-072A-50
16	371.5	19.2	353	3	US-08-986-485-6
17	369.5	19.1	1529	4	US-09-312-283C-396
18	359.5	18.6	1525	3	US-09-191-647-2
19	359.5	18.6	1525	3	US-09-540-245A-2
20	359.5	18.6	1525	3	US-09-540-153-2
21	357	18.4	1523	3	US-09-182-024A-2
22	339	17.5	907	4	US-09-170-496D-264
23	339	17.5	907	4	US-09-170-496D-278
24	333.5	17.2	1119	4	US-09-907-794A-294
25	333.5	17.2	1119	4	US-09-905-125A-294
26	333.5	17.2	1119	4	US-09-902-775A-294
27	333.5	17.2	1119	4	US-09-906-700-294

28	333.5	17.2	1119	4	US-10-140-002-352	Sequence 352, App
29	333.5	17.2	1119	4	US-09-903-603A-294	Sequence 294, App
30	327	16.9	1091	3	US-08-986-485-5	Sequence 5, Appli
31	323.5	16.7	642	4	US-10-140-002-370	Sequence 370, App
32	321.5	16.6	620	4	US-09-907-794A-73	Sequence 73, Appl
33	321.5	16.6	620	4	US-09-905-125A-73	Sequence 73, Appl
34	321.5	16.6	620	4	US-09-902-775A-73	Sequence 73, Appl
35	321.5	16.6	620	4	US-09-906-700-73	Sequence 73, Appl
36	321.5	16.6	620	4	US-10-140-002-324	Sequence 324, App
37	321.5	16.6	620	4	US-09-903-603A-73	Sequence 73, Appl
38	319.5	16.5	1480	3	US-09-191-647-7	Sequence 7, Appli
39	319.5	16.5	1480	3	US-09-540-245A-7	Sequence 7, Appli
40	319.5	16.5	1480	3	US-09-540-153-7	Sequence 7, Appli
41	319.5	16.5	1480	3	US-09-182-024A-5	Sequence 5, Appli
42	319.5	16.5	1480	5	PCT-US91-09055-2	Sequence 2, Appli
43	313.5	16.2	708	3	US-09-131-648-2	Sequence 2, Appli
44	313.5	16.2	708	4	US-09-907-794A-69	Sequence 69, Appl
45	313.5	16.2	708	4	US-09-905-125A-69	Sequence 69, Appl

ALIGNMENTS

RESULT 1  
US-08-592-500-2  
; Sequence 2, Application US/08592500  
; Patent No. 6005089  
; GENERAL INFORMATION:  
; APPLICANT: Lanza, Francois  
; APPLICANT: Phillips, David R.  
; APPLICANT: Cazenave, Jean-Pierre  
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/592,500  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/089,455  
; FILING DATE: 09-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,684  
; REFERENCE/DOCKET NUMBER: 12418-28  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 560 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-592-500-2

Query Match 76.2%; Score 1477; DB 3; Length 560;  
Best Local Similarity 76.1%; Pred. No. 1.9e-135;  
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

Oy 1 MLRSALLSAVLALRAQPPFPCPKTKCVVRDAACSGSVAHIAELGLPTNLTHLLFRM 60

Db 1 MLRGTLLCAVLGLLRAQPPPCPKCKCVFRDAAQCSCGGDVARIASALGLPTNLTHILLFGM 60  
QY 61 DOGILRNHSFGMTVLQRLMLSDSHISALDPGTFTNDLVKLTIRLNRKISRPRAILDK 120  
Db 61 GRGVLSQSFSGMTVLQRLMISDSHISAVAPGTFTSDLIKLTIRLSRNKITHLPALLDK 120  
QY 121 MVLEQLFLDHNAALRDLDQNLFOQLRNQLQELGLNQNLSPANLFSLSRELKLDLSRN 180  
Db 121 MVLEQLFLDHNAALRGIDQNMFKLVNLQELALNQNLDFLPASLFTNLENLKLDLSGN 180  
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRNLERNHLSRVAPGAFDR 240  
Db 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRNLERNHLSRVAPGAFDR 240  
QY 241 LGNLSLTLSCNLLIESLPPALFLHVSVSRLTLFENPLEELPDVLFGEAGRELWLNLT 300  
Db 241 LPNLSSLTSLRNHLAFLPSALFLHSHNLTLTLFENPLAELPGVLFGEAGRELWLNLT 300  
QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRISALPRGVFQGLRELRLVLAHTNALAEIRDDALR 360  
Db 301 QLRTLPAAAFRNLSRLRYLGVTLSPRLSALPQGAFOGLGELQVLAHNSGLTALPDGLLR 360  
QY 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389  
Db 361 GLGKLQVSLRNHLRALPRLFRNLSSL 389

RESULT 2

US-08-195-006-2  
; Sequence 2, Application US/08195006  
; Patent No. 6083688  
; GENERAL INFORMATION:  
; APPLICANT: Lanza, Francois  
; APPLICANT: Phillips, David R.  
; APPLICANT: Cazenave, Jean-Pierre  
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/195,006  
; FILING DATE: 10-FEB-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/089,455  
; FILING DATE: 09-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,684  
; REFERENCE/DOCKET NUMBER: 12418-28  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 560 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-195-006-2

Query Match 76.2%; Score 1477; DB 3; Length 560;  
Best Local Similarity 76.1%; Pred. No. 1.9e-135;

Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;  
QY 1 MLRSALLSAVLALRAQPPPCPKCKCVVRDAAQCSCGGSVAHIAELGLPTNLTHILLFRM 60  
Db 1 MLRGTLLCAVLGLLRAQPPPCPKCKCVFRDAAQCSCGGDVARIASALGLPTNLTHILLFGM 60  
QY 61 DOGILRNHSFGMTVLQRLMLSDSHISALDPGTFTNDLVKLTIRLNRKISRPRAILDK 120  
Db 61 GRGVLSQSFSGMTVLQRLMISDSHISAVAPGTFTSDLIKLTIRLSRNKITHLPALLDK 120  
QY 121 MVLEQLFLDHNAALRDLDQNLFOQLRNQLQELGLNQNLSPANLFSLSRELKLDLSRN 180  
Db 121 MVLEQLFLDHNAALRGIDQNMFKLVNLQELALNQNLDFLPASLFTNLENLKLDLSGN 180  
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRNLERNHLSRVAPGAFDR 240  
Db 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRNLERNHLSRVAPGAFDR 240  
QY 241 LGNLSLTLSCNLLIESLPPALFLHVSVSRLTLFENPLEELPDVLFGEAGRELWLNLT 300  
Db 241 LPNLSSLTSLRNHLAFLPSALFLHSHNLTLTLFENPLAELPGVLFGEAGRELWLNLT 300  
QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRISALPRGVFQGLRELRLVLAHTNALAEIRDDALR 360  
Db 301 QLRTLPAAAFRNLSRLRYLGVTLSPRLSALPQGAFOGLGELQVLAHNSGLTALPDGLLR 360  
QY 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389  
Db 361 GLGKLQVSLRNHLRALPRLFRNLSSL 389

RESULT 3

US-09-063-950-4  
; Sequence 4, Application US/09063950C  
; Patent No. 6225085  
; GENERAL INFORMATION:  
; APPLICANT: Holtzman, Douglas A.  
; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: MEI-019  
; CURRENT APPLICATION NUMBER: US/09/063,950C  
; CURRENT FILING DATE: 1998-04-21  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent in ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-063-950-4

Query Match 76.2%; Score 1477; DB 3; Length 560;  
Best Local Similarity 76.1%; Pred. No. 1.9e-135;

Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;  
QY 1 MLRSALLSAVLALRAQPPPCPKCKCVVRDAAQCSCGGSVAHIAELGLPTNLTHILLFRM 60  
Db 1 MLRGTLLCAVLGLLRAQPPPCPKCKCVFRDAAQCSCGGDVARIASALGLPTNLTHILLFGM 60  
QY 61 DOGILRNHSFGMTVLQRLMLSDSHISALDPGTFTNDLVKLTIRLNRKISRPRAILDK 120  
Db 61 GRGVLSQSFSGMTVLQRLMISDSHISAVAPGTFTSDLIKLTIRLSRNKITHLPALLDK 120  
QY 121 MVLEQLFLDHNAALRDLDQNLFOQLRNQLQELGLNQNLSPANLFSLSRELKLDLSRN 180  
Db 121 MVLEQLFLDHNAALRGIDQNMFKLVNLQELALNQNLDFLPASLFTNLENLKLDLSGN 180  
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRNLERNHLSRVAPGAFDR 240  
Db 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRNLERNHLSRVAPGAFDR 240  
QY 241 LGNLSLTLSCNLLIESLPPALFLHVSVSRLTLFENPLEELPDVLFGEAGRELWLNLT 300  
Db 241 LPNLSSLTSLRNHLAFLPSALFLHSHNLTLTLFENPLAELPGVLFGEAGRELWLNLT 300



QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAELEDDALR 360  
Db 301 QLRTLPAAAFRNLSRLRYLGVTLSPRLSALPQGAFOGLGELQVLALHSNGLTALPDGLLR 360  
QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389  
Db 361 GLGKLRQVSLRRNRLRALPRLFRNLSSL 389

RESULT 4  
PCT-US94-07644A-2  
; Sequence 2, Application PC/TUS9407644A  
; GENERAL INFORMATION:  
; APPLICANT: COR Therapeutics, Inc.  
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/07644A  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,684  
; REFERENCE/DOCKET NUMBER: 012418-003000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 560 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-07644A-2

Query Match 76.2%; Score 1477; DB 5; Length 560;  
Best Local Similarity 76.1%; Pred. No. 1.9e-135;  
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;  
QY 1 MLRSALLSAVALLRAQFPCKTKCKVVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60  
Db 1 MLRGILLCAVLGLLRAQFPCKCKVFRDAAQCSGSDVARISALGLPTNLTHILLFGM 60  
QY 61 DQGLRNHFSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLKTLRLTRNKISRLPRAILDK 120  
Db 61 GRGVLSQSFSFGMTVLQRLMLSDSHISAVAPGTFSDLIKLTLSRNKITHLPGALLDK 120  
QY 121 MVLLQLFLDHNAIRDLDQNLFPQQLRNQLGELGNONQLSFLPANLFSLSRELKLLDLSRN 180  
Db 121 MVLLQLFLDHNAIRGIDQNMFKLVNQLQELALNQNQLDFLPASLFTNLNKLDDLSGN 180  
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNGLALTELRLERNHLSRVAPGAFDR 240  
Db 181 NLTHLPKGLLGAQAKLERLLHSNRLVSLDGLLSNGLALTELQFHRNHRSIAPGAFDR 240  
QY 241 LGNLSLTLTSGNLLSPALFLHVSSVSRLLTFENPLEELPDVLFGEAGLRELWLNGT 300  
Db 241 LPNLSLTLSRNHLAFLPSALFLHSNLTLLTFENPLAELPGVLFGEAGLQELWLNRT 300  
QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAELEDDALR 360

Db 301 QLRTLPAAAFRNLSRLRYLGVTLSPRLSALPQGAFOGLGELQVLALHSNGLTALPDGLLR 360  
QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389  
Db 361 GLGKLRQVSLRRNRLRALPRLFRNLSSL 389  
RESULT 5  
US-08-190-802A-49  
; Sequence 49, Application US/08190802A  
; Patent No. 5519003  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: P.O. Box 60850  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-0850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/190,802A  
; FILING DATE: 01-FEB-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: 8600-0139  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 49:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 605 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: Insulin-like growth factor binding  
; INDIVIDUAL ISOLATE: protein complex, Fig. 32  
US-08-190-802A-49

Query Match 21.6%; Score 419.5; DB 1; Length 605;  
Best Local Similarity 33.0%; Pred. No. 3.7e-32;  
Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;  
QY 38 GSAH---IAELGLPTNLTHILLFRMDQGLRNHFSFGMTVLQRLMLSDSHISAIDPGTF 94  
Db 140 GTFATFPALASGLSNN---RLSRLEDGL-----FEGLSLWDLNLGWSLAVLPDAAF 190  
QY 95 NDLVKLTLRLTRNKISRLPRAILDKMVLEQLFLDHNAIRDLDQNLFPQQLRNQLGELGN 154  
Db 191 RGLSRLRELVLAGNRLAYLPALFSGLAELRELDLSRNALRAIKANVFQVLPRLQKLYLD 250  
QY 155 QNQLSFLPANLFSLSRELKLLDLSRNNLTHLPK----GLLGAQVKLEKLLYSNQLTSVD 210  
Db 251 RNLIAAVAPGAFGLKALRWLDLSHNRVAGLLEDTFGLLGLRV---LRLSHNAIASLR 306  
QY 211 SGLLSNGLALTELRLERNHLSRVAPGAFDRGLNLSLITLSGNLLSPALFLHVSSVSR 270  
Db 307 PRTEKDLHFEELQLGHNRIQLAERSFEGLGQLEVLTDHNLQLEVKAGAFGLGNTNAV 366



Db 140 GTFHTPALASGLSNN-----RLSRLEDGL-----FEGLSLWDNLNGWSLAVLPDAAF 190  
QY 95 NDLVKLKTRLTRNKISRPAILDKMVLLQQLFBDHNAIRDLDQNLFOQLRNLOELGLN 154  
Db 191 RGLSLRELVLAGNRLAYLOPALFSGLAELRELDLSRNALRAIKANVFVQLPRLQKLYLD 250  
QY 155 QNQLSFLPANLFSLSRELKLLDLSRNNLTHLPK-----GLLGAQVKLEKLLLYSNQLTSDV 210  
Db 251 RNLIAAVAPGAFGLGKALRWLDLSHNRVAGLLEDTFPGLLGLRV---LRLSHNAIASLR 306  
QY 211 SGLLSNLGALTELRLERNHLRSVAPGAFDRIGNLSSLTSGNLLLESPPALFLHVSSVSR 270  
Db 307 PRTEKDLHFLEELQGHNRIRQLAERSFEGLGQLEVLTLTDHNLQLEVKAGAFGLGTNVAV 366  
QY 271 LTLFENPLEELPDVLFGEAGRELWNGTHLSTLPAAAFRNLSGLQTLGLTRN----- 324  
Db 367 MNLGNCRLNLPQOVFRGLGKLSLHLEGSCLGRIRPHTFTGLSGLRRLFLKDNGLVGIE 426  
QY 325 -----PRLSALPRGVFQGLRELRLVLAHTNALAEIRDALRGLGHLRQ 367  
Db 427 EQSLWGIAELLELDLTSNQLTHLPHRLFQGLGKLEYLLLSRNRLAELPADALGPLQRAFW 486  
QY 368 VSLRHNRLRALPRTLFRNLSSL 389  
Db 487 LDVSHNRLEALPNSLLAPLGR 508

RESULT 8  
US-08-487-072A-49  
; Sequence 49, Application US/08487072A  
; Patent No. 6423684  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,072A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 2550-0025.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763  
; INFORMATION FOR SEQ ID NO: 49:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 605 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: Insulin-like growth factor binding  
; INDIVIDUAL ISOLATE: protein complex, Fig. 32  
US-08-487-072A-49

Query Match 21.6%; Score 419.5; DB 4; Length 605;  
Best Local Similarity 33.0%; Pred. No. 3.7e-32;  
Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;  
QY 38 GSAVAH---IAELGLPTNLTLLFRMDQGIILRNHFSGMTVLQRLMLSDSHISAIDPGTF 94  
Db 140 GTFHTPALASGLSNN-----RLSRLEDGL-----FEGLSLWDNLNGWSLAVLPDAAF 190  
QY 95 NDLVKLKTRLTRNKISRPAILDKMVLLQQLFBDHNAIRDLDQNLFOQLRNLOELGLN 154  
Db 191 RGLSLRELVLAGNRLAYLOPALFSGLAELRELDLSRNALRAIKANVFVQLPRLQKLYLD 250  
QY 155 QNQLSFLPANLFSLSRELKLLDLSRNNLTHLPK-----GLLGAQVKLEKLLLYSNQLTSDV 210  
Db 251 RNLIAAVAPGAFGLGKALRWLDLSHNRVAGLLEDTFPGLLGLRV---LRLSHNAIASLR 306  
QY 211 SGLLSNLGALTELRLERNHLRSVAPGAFDRIGNLSSLTSGNLLLESPPALFLHVSSVSR 270  
Db 307 PRTEKDLHFLEELQGHNRIRQLAERSFEGLGQLEVLTLTDHNLQLEVKAGAFGLGTNVAV 366  
QY 271 LTLFENPLEELPDVLFGEAGRELWNGTHLSTLPAAAFRNLSGLQTLGLTRN----- 324  
Db 367 MNLGNCRLNLPQOVFRGLGKLSLHLEGSCLGRIRPHTFTGLSGLRRLFLKDNGLVGIE 426  
QY 325 -----PRLSALPRGVFQGLRELRLVLAHTNALAEIRDALRGLGHLRQ 367  
Db 427 EQSLWGIAELLELDLTSNQLTHLPHRLFQGLGKLEYLLLSRNRLAELPADALGPLQRAFW 486  
QY 368 VSLRHNRLRALPRTLFRNLSSL 389  
Db 487 LDVSHNRLEALPNSLLAPLGR 508

RESULT 9  
US-09-538-092-1087  
; Sequence 1087, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CurapatSeqFormatter Version 0.9  
; SEQ ID NO 1087  
; LENGTH: 605  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (0)..(0)  
; OTHER INFORMATION: Polypeptide Accession Number P35858  
US-09-538-092-1087

Query Match 21.6%; Score 419.5; DB 4; Length 605;  
Best Local Similarity 33.0%; Pred. No. 3.7e-32;  
Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;  
QY 38 GSAVAH---IAELGLPTNLTLLFRMDQGIILRNHFSGMTVLQRLMLSDSHISAIDPGTF 94  
Db 140 GTFHTPALASGLSNN-----RLSRLEDGL-----FEGLSLWDNLNGWSLAVLPDAAF 190  
QY 95 NDLVKLKTRLTRNKISRPAILDKMVLLQQLFBDHNAIRDLDQNLFOQLRNLOELGLN 154  
Db 191 RGLSLRELVLAGNRLAYLOPALFSGLAELRELDLSRNALRAIKANVFVQLPRLQKLYLD 250





; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-0850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/190,802A  
; FILING DATE: 01-FEB-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: 8600-0139  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 603 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: Insulin-like growth factor bind.  
; INDIVIDUAL ISOLATE: pro. complex-rat, Fig. 33  
; US-08-190-802A-50

Query Match 19.7%; Score 382.5; DB 1; Length 603;  
Best Local Similarity 28.7%; Pred. No. 1.5e-28;  
Matches 121; Conservative 64; Mismatches 182; Indels 55; Gaps 7;

QY 21 CPKCKCVVRD-----AAQSGGSAHIAELGLPTNLTHILLFRMDQGI LRNHSFGMTV 75  
Db 41 CPVACTCSHDDYTDLSVFCSSKNLTHLPD-DIPVSTRALWLDGNNLSSIPSAFQNLSS 99  
QY 76 LQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKISRLPRAILDKMVLLLEQLFLDHNALR 135  
Db 100 LDFNLQGSWLRSLPEQALLGLQNLYYLHLERNRLRLNLA VGLFTHTPTSLASLSLSSNLLG 159  
QY 136 LLDQNLFQQLRNQLQELGNQNLSPFLPANLFSLSRELKLLDLSRNNLTHLPKGLLGAQVK 195  
Db 160 RLEEGLFQGLSHLWDLNLGWSLVVLPDPTVFQGLGNLHELVL AGNKLTYLPALFCGLGE 219  
QY 196 LEKLLYSNQLTSVDSGLLSNLGALTELRLERNHLSRVAPGAF-----DRLG 242  
Db 220 LRELDLSRNALRSVKANFVHLPRQKLYLDRNLITAVAPGAF LGMKALRWLDLSHNRVA 279  
QY 243 NLSS-----LTLSGNLLSPPALFLHVSSVSRLTLFENPLEELPDVLFGE MAG 291  
Db 280 GLMEDTFPGLLGLHVLRLAHNAIASLRPRTFKDLHLEELQ LGHNRIRQLGERTFEGLGQ 339  
QY 292 LRELWNGTHLSTLPAAAFRLNSGLQTLGLTRNPRLSALPRGV FQGLRELRLV----- 344  
Db 340 LEVLTLDNDNQITEVRVGAFSGLFNVAVMNLSGN-CLRS LPERVFGQLDKLSLHLSHLSCL 398  
QY 345 ---ALHT-----NALAELRDDALRGHLRQVSLRHNRLRALPRTLFRNLS 387  
Db 399 GHVRLHTFAGLSGLRRLFLRDNSSISIEEQSLAGLSLELDELDTTNR LTHLPRLQFLQGLG 458  
QY 388 SL 389  
Db 459 HL 460

RESULT 13  
US-08-477-346-50  
; Sequence 50, Application US/08477346

; Patent No. 6262023  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Daria  
; APPLICANT: Ron, Dorit  
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
; TITLE OF INVENTION: Thereof  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 2000 Pennsylvania Avenue, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,346  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/487,072  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 2550-0025.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 603 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: Insulin-like growth factor bind.  
; INDIVIDUAL ISOLATE: pro. complex-rat, Fig. 33  
; US-08-477-346-50

Query Match 19.7%; Score 382.5; DB 3; Length 603;  
Best Local Similarity 28.7%; Pred. No. 1.5e-28;  
Matches 121; Conservative 64; Mismatches 182; Indels 55; Gaps 7;

QY 21 CPKCKCVVRD-----AAQSGGSAHIAELGLPTNLTHILLFRMDQGI LRNHSFGMTV 75  
Db 41 CPVACTCSHDDYTDLSVFCSSKNLTHLPD-DIPVSTRALWLDGNNLSSIPSAFQNLSS 99  
QY 76 LQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKISRLPRAILDKMVLLLEQLFLDHNALR 135  
Db 100 LDFNLQGSWLRSLPEQALLGLQNLYYLHLERNRLRLNLA VGLFTHTPTSLASLSLSSNLLG 159  
QY 136 LLDQNLFQQLRNQLQELGNQNLSPFLPANLFSLSRELKLLDLSRNNLTHLPKGLLGAQVK 195  
Db 160 RLEEGLFQGLSHLWDLNLGWSLVVLPDPTVFQGLGNLHELVL AGNKLTYLPALFCGLGE 219  
QY 196 LEKLLYSNQLTSVDSGLLSNLGALTELRLERNHLSRVAPGAF-----DRLG 242  
Db 220 LRELDLSRNALRSVKANFVHLPRQKLYLDRNLITAVAPGAF LGMKALRWLDLSHNRVA 279  
QY 243 NLSS-----LTLSGNLLSPPALFLHVSSVSRLTLFENPLEELPDVLFGE MAG 291  
Db 280 GLMEDTFPGLLGLHVLRLAHNAIASLRPRTFKDLHLEELQ LGHNRIRQLGERTFEGLGQ 339  
QY 292 LRELWNGTHLSTLPAAAFRLNSGLQTLGLTRNPRLSALPRGV FQGLRELRLV----- 344  
Db 340 LEVLTLDNDNQITEVRVGAFSGLFNVAVMNLSGN-CLRS LPERVFGQLDKLSLHLSHLSCL 398





***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 13:05:08 ; Search time 25 Seconds  
(without alignments)  
1497.133 Million cell updates/sec

Title: US-09-775-803-12\_COPY\_1\_389  
Perfect score: 1938  
Sequence: 1 MLRSALLSAVLALLRAQPPF.....LRHNRDLRALPRTLFRNLSSL 389

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %			DB ID	Description
	Score	Match	Length		
1	1477	76.2	560	2 A60164	platelet membrane
2	419.5	21.6	605	2 A41915	insulin-like growth
3	416.5	21.5	536	2 A34901	lysine carboxypept
4	413.5	21.3	605	2 JC5239	insulin-like growth
5	382.5	19.7	603	2 JC1282	insulin-like growth
6	382	19.7	603	2 JC6128	insulin-like growth
7	379.5	19.6	1531	2 T42218	slit-1 protein hom
8	367.5	19.0	907	2 JG0193	G protein-coupled
9	366	18.9	1523	2 T13953	MEGF5 protein - ra
10	345	17.8	1389	2 T13852	Gene wheeler prote
11	339	17.5	907	2 JE0176	Orphan G protein-c
12	337	17.4	1051	2 T13174	gp150 protein - fr
13	337	17.4	1385	2 T13887	tlr protein - frui
14	327	16.9	1091	2 A58532	glial cell membran
15	319.5	16.5	1469	2 B36665	slit protein 2 pre
16	319.5	16.5	1480	2 A36665	slit protein 1 pre
17	317.5	16.4	707	2 JC7763	neuronal leucine-r
18	317	16.4	361	2 A53860	chondroadherin pre
19	313	16.2	1066	2 T15864	hypothetical prote
20	311.5	16.1	622	2 JC7973	synleucin - human
21	310	16.0	575	2 T29972	hypothetical prote
22	306.5	15.8	1134	1 A29944	choaptin precursor
23	305.5	15.8	1097	2 A29943	Toll protein precu
24	305	15.7	961	2 T23395	hypothetical prote
25	302	15.6	610	2 T23836	hypothetical prote
26	301	15.5	594	2 T23841	hypothetical prote
27	300	15.5	789	2 T28714	hypothetical prote
28	300	15.5	1355	2 T28715	hypothetical prote
29	298.5	15.4	1119	2 AD1822	leucine-rich-repea

30	296	15.3	662	2 S42799	garp precursor - h
31	295.5	15.2	1495	2 T31434	densin-180 - rat
32	291	15.0	680	2 T19939	hypothetical prote
33	290.5	15.0	738	2 T19938	hypothetical prote
34	288.5	14.9	1112	2 T10504	disease resistance
35	286.5	14.8	572	2 T30947	hypothetical prote
36	286.5	14.8	613	2 A88684	protein AC7.2 (imp
37	286	14.8	526	2 C84552	hypothetical prote
38	283.5	14.6	312	1 NBHUA2	hypothetical prote
39	283.5	14.6	559	2 T42998	leucine-rich alpha
40	269	13.9	682	2 A49121	Ras-binding protei
41	269	13.9	682	2 A43318	cell-surface molec
42	266	13.7	1256	2 S60461	connectin precurs
43	265.5	13.7	1016	2 T30553	gene flightless-I
44	264.5	13.6	333	2 T34555	disease resistance
45	253	13.1	1232	2 T05322	hypothetical prote

ALIGNMENTS

RESULT 1

A60164  
platelet membrane glycoprotein V precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 12-Jan-1993 #sequence\_revision 24-Feb-1994 #text\_change 09-Jul-2004  
C;Accession: A48030; A60164; A35483; B35483; C35483; A60432; A47507; S34329  
R;Lanza, F.; Morales, M.; de La Salle, C.; Cazenave, J.P.; Clemetson, K.J.; Shimomura, T  
J. Biol. Chem. 268, 20801-20807, 1993  
A;Title: Cloning and characterization of the gene encoding the human platelet glycoprote  
A;Reference number: A48030; MUID:94012616; PMID:8407908  
A;Accession: A48030  
A;Molecule type: DNA  
A;Residues: 1-560 <LA2>  
A;Cross-references: UNIPROT:P40197; EMBL:Z23091; NID:g312501; PIDN:CAA80637.1; PID:g3125  
R;Shimomura, T.; Fujimura, K.; Maehama, S.; Takemoto, M.; Oda, K.; Fujimoto, T.; Oyama.  
Blood 75, 2349-2356, 1990  
A;Title: Rapid purification and characterization of human platelet glycoprotein V: the a  
A;Reference number: A60164; MUID:90275263; PMID:2350580  
A;Accession: A60164  
A;Molecule type: protein  
A;Residues: 365-384,'X',386-390,'X',392-395,'X',397;188-208,'I',210;27-50,'X',52-53;174-  
, 'XX',108,'T',61-72,'TK',75-77,'V',56-57,'G',479-487,'X',489-498,'X',500,'X',502-503,'X',  
R;Roth, G.J.; Church, T.A.; McMullen, B.A.; Williams, S.A.  
Biochem. Biophys. Res. Commun. 170, 153-161, 1990  
A;Title: Human platelet glycoprotein V: a surface leucine-rich glycoprotein related to  
A;Reference number: A35483; MUID:90321220; PMID:2372284  
A;Accession: A35483  
A;Molecule type: protein  
A;Residues: 145-166,'I',168-169,'X',171-172 <ROT>  
A;Note: this proteolytic fragment was designated peptide M392  
A;Accession: B35483  
A;Molecule type: protein  
A;Residues: 121-129,'W',131-135;466-468,'X',470 <RO2>  
A;Note: this material was designated peptide M393 but may contain two peptides  
A;Accession: C35483  
A;Molecule type: protein  
A;Residues: 252-266,'H',268-272,'X',274-279,'I',281-284,'I',286 <RO3>  
A;Note: this proteolytic fragment was designated peptide M401  
R;Zafar, R.S.; Walz, D.A.  
Thromb. Res. 53, 31-44, 1989  
A;Title: Platelet membrane glycoprotein V: characterization of the thrombin-sensitive gl  
A;Reference number: A60432; MUID:89162331; PMID:2922700  
A;Accession: A60432  
A;Molecule type: protein  
A;Residues: 477-478,'FX',481-485,'E',487,'V',489-492,'NQ',495,'E',497-498 <ZAF>  
R;Hickey, M.J.; Hagen, F.S.; Yagi, M.; Roth, G.J.  
Proc. Natl. Acad. Sci. U.S.A. 90, 8327-8331, 1993  
A;Title: Human platelet glycoprotein V: characterization of the polypeptide and the rel.  
A;Reference number: A47507; MUID:93391348; PMID:7690959  
A;Accession: A47507  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA

A:Residues: 1-560 <RES>  
A:Cross-references: GB:L11238; NID:g388759; PIDN:AAA03069.1; PID:g388760  
C:Comment: This platelet membrane protein is a substrate for thrombin.  
C:Comment: The amino end of the intact protein is blocked.  
C:Comment: This protein is absent in Bernard-Soulier syndrome.  
C:Genetics:  
A:Gene: GDB:GP5  
A:Cross-references: GDB:230236; OMIM:173511  
A:Map position: 5pter-5qter  
C:Keywords: blocked amino end; glycoprotein; platelet; tandem repeat; transmembrane protein

RESULT 2

A41915

insulin-like growth factor-binding complex acid-labile chain precursor - human

N:Alternate names: Acid-Labile Subunit (ALS)

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004

C:Accession: A41915

R:Leong, S.R.; Baxter, R.C.; Camerato, T.; Dai, J.; Wood, W.I.

Mol. Endocrinol. 6, 870-876, 1992

A:Title: Structure and functional expression of the acid-labile subunit of the insulin-L

A:Reference number: A41915; MUID:92357025; PMID:1379671

A:Accession: A41915

A:Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 1-605 <LEQ>

A:Cross-references: UNIPROT:P35858; GB:M36826; NID:G184807; PIDN:AAA35047.1; PID:G184808

A:Experimental source: liver

A:Note: sequence extracted from NCBI backbone (NCBIP:110171)

F:75-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:99-122/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:123-146/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:147-170/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:171-194/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:195-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:219-242/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F:243-266/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>

F:267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>

F:291-314/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>

F:315-338/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F;339-362/Domain:	leucine-rich	alpha-2-glycoprotein	repeat	homology	<LR12>
F;363-386/Domain:	leucine-rich	alpha-2-glycoprotein	repeat	homology	<LR13>
F;387-410/Domain:	leucine-rich	alpha-2-glycoprotein	repeat	homology	<LR14>
F;411-434/Domain:	leucine-rich	alpha-2-glycoprotein	repeat	homology	<LR15>
F;435-458/Domain:	leucine-rich	alpha-2-glycoprotein	repeat	homology	<LR16>
F;459-482/Domain:	leucine-rich	alpha-2-glycoprotein	repeat	homology	<LR17>
F;483-506/Domain:	leucine-rich	alpha-2-glycoprotein	repeat	homology	<LR18>
F;507-529/Domain:	leucine-rich	alpha-2-glycoprotein	repeat	homology	<LR19>

  

Query Match	21.6%;	Score	419.5;	DB 2;	Length	605;			
Best Local Similarity	33.0%;	Pred.	No. 2.1e-21;						
Matches	126;	Conservative	58;	Mismatches	155;	Indels	43;	Gaps	6

  

QY	38	GSVAH---	IAELGLPTNLTHILLFRMDQGILRNHPSGMTVQLRRLMLSDSHISAIDPGTF	94
DB	140	GTFAHTPALASGLSNN---	RLSRLEDGL-----FEGLSGLWDLNLGWSLAVLPDAAF	190
QY	95	NDLVKIKTLRLTRNKISRLPRAILDKMWLLEQLFLDHNAIRDLDQNLFOQLRNLOELGLN	154	
DB	191	RGIGSLRELVLAGNRLAYLPALFSGLAELRELDLSRNALRAIKANVFQVLPRLQKLYLD	250	
QY	155	QNQLSFELPANLFSSLRELKLLDLSRNNLTHLPK---	GLLGAQVKLEKLLLYSNQLTSVD	210
DB	251	RNLIAAVAPGAFGLKALRWDLDSHNRVAGLLEDFPGLLGLRV---	LRLSHNAIASLR	306
QY	211	SGLLSNIGALTELRLERNHLRSVAPGAFDRGLGNLSLTLTSGNLLLESPPALFLHVSSVSR	270	
DB	307	PRTFKDLHFLEELQLGHNRIQLAERSFEGLGQLEVLTLTDHNLQLEVKACAFGLGTNVAV	366	
QY	271	LTLFENPLEELPDVLFEGMAGLRELWLNGTHLSLTPAAAFRNLSGLQTLGLTRN-----	324	
DB	367	MNLSGNCLRNLPQOVFRGLKHLHLHLEGGSCIGRIPIPHFTFTGLSGRLRLFKDNGLVGIE	426	
QY	325	-----PRISALPRGVFQGLRELRVLALHTNALAELRDDALRGGLHLRQ	367	
DB	427	EQSLWGLAEELLELDLTSNQLTHLPRLFOGLGKLEYLLLSRNRLAELPADALGPLQRAFW	486	
QY	368	VSLRRNHLRALPRTLFRNLSSL	389	
DB	487	LDVSHNELEALPNLSLLAPLGL	508	

RESULT 3  
A34901  
lysine carboxypeptidase (EC 3.4.17.3) 83K chain - human  
C;Species: Homo sapiens (man)  
C;Date: 20-Jul-1990 #sequence\_revision 20-Jul-1990 #text\_change 09-Jul-2004  
C;Accession: A34901  
R;Tan, F.; Weerasinghe, D.K.; Skidgel, R.A.; Tamei, H.; Kaul, R.K.; Roninson, I.B.; Schijl  
J. Biol. Chem. 265, 13-19, 1990  
A;Title: The deduced protein sequence of the human carboxypeptidase N high molecular wei  
A;Reference number: A34901; MUID:90094386; PMID:2378615  
A;Accession: A34901  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-536 <TAN>  
A;Cross-references: UNIPROT:P22792; GB:J05158; NID:gl79935; PIDN:AAAS1921.1; PID:gl79936  
C;Genetics:

F:341-364/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>	
Query Match	21.5%; Score 416.5; DB 2; Length 536;
Best Local Similarity	32.3%; Pred. No. 2.9e-21;
Matches	115; Conservative 53; Mismatches 163; Indels 25; Gaps 3;
QY	22 PKTKCKVVRDAAQCSCGGVAHI-----AELGPTNLTHILLF 58
Db	2 PMGDCFCVQEVF-CSDEELATVPLDIPPYTKNIIFVETSFTTLETRAFGSNPNLTKWFL 60
QY	59 RNDQGILRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTLRTRNKISRLPRAIL 118
Db	61 DTQLCQFRPDAGGLPRLEDEVTGSSFLNLSTNIFSNTLSGLKLTNFNMLEALPEGLF 120
QY	119 DKMVLLEQLFDHNALRDQNLFQQLRNLQELGLNQQLSFLPANLFSLSRELKLLDLS 178
Db	121 QHLAALESHLQGNQLOALPRRLFQPLTHLKTNLNAQNLLAQLPEELFHPHLSLQTLKLS 180
QY	179 RNNLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRLERNHLRSVAPGAF 238
Db	181 NNALSGLPQGVFGKLSQLBELFDLSNNISELPPQVFSQLFCLERLWLQRNAITHLPLSIF 240
QY	239 DRLGNLSLTLSGNLLESIPPALFLHVSSVSRLLTFENPLEELPDVLFGEWAGLRELWLN 298
Db	241 ASLGNLTFLSLQWNMLRVLPAGLFAHTPCVLGLSLTHNQLETVTGTFHAHLSNLSLMLS 300
QY	299 GTHLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAEL 354
Db	301 YNAITHLPAGIFRDLEELVKLYLGSN-NLTALHFAHFQNLKLELLSLSKNQLTTL 355
RESULT 4	
JC5239	
insulin-like growth factor acid-labile chain - baboon	
C:Species: Papio sp. (baboon)	
C:Date: 17-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-May-1997	
C:Accession: JC5239	
R;Delhanty, P.; Baxter, R.C.	
Biochem. Biophys. Res. Commun. 227, 897-902, 1996	
A:Title: The cloning and expression of the baboon acid-labile subunit of the insulin-like	
A:Reference number: JC5239; MUID:97040714; PMID:8886027	
A:Contents: liver	
A:Accession: JC5239	
A:Molecule type: mRNA	
A:Residues: 1-605 <DEL>	
C:Comment: This factor is structurally related to proinsulin and have insuline-like meta	
Query Match	21.3%; Score 413.5; DB 2; Length 605;
Best Local Similarity	29.8%; Pred. No. 5.4e-21;
Matches	141; Conservative 56; Mismatches 167; Indels 109; Gaps 8;
QY	21 CPKTKCKVVRDAAQ-----CSG----- 37
Db	41 CPATCACSYDDEVNELSVFCSSRNLTRPDGIPGGTQALWLDSDNNLSSIPPAAFRNLSL 100
QY	38 -----GSVAHIAELGLPTNLTHILLFRMDQGILRNHSFGMTVLQRLMLSDSHIS 87
Db	101 AFLNLQGGQLGSLEPQALLGL-ENLCHLHLERNQLRS LAVGTFAYTPALALLGLSNNRLS 159
QY	88 AIDPGTFNDLVKLTIRLRNKISRLPRAI-----LDKMVL----- 123
Db	160 RLEDGLFEGGLGNLWDLNLGWNSLAVLPDAAFRGLGGLRELVLAGNRLAYLQPALFSGLAE 219
QY	124 LEQLFLDHNALRDQNLFQQLRNLQELGLNQQLSFLPANLFSLSRELKLLDLSRNNLT 183
Db	220 LRELDLSRNALRAIKANVFAQLPRQLKLYLDRNLIAAVAPGAFGLKALRWLDLSHNRVA 279
QY	184 HLPK----GLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRLERNHLRSVAPGAFD 239
Db	280 GLEEDTFPGLGLGRV----LRLSHNAIASLRPTFEDLHFLEELQLGHNRIRQLAERSFE 335
QY	240 RLGNLSLTLSGNLLESIPPALFLHVSSVSRLLTFENPLEELPDVLFGEWAGLRELWLNG 299

Db	336 GLGQLEVLTLDHNQLEQVKVGAFLGLTNVAVMNLSGNCLRNLP EQVFRGLGKLHSLHLEG 395
QY	300 THLSTLPAAAFRNLSGLQTLGLTRN-----PRLSALPRGVFQ 336
Db	396 SCLGRIRPHTFAGLSGLRRFLKDNGLVGEIEEQSLWGLABELDLTNSQLTHLPQLFQ 455
QY	337 GLRELRLVALHTNALAELRDDALRGHLRQVSLRHNRLRALPRTLFRNLSL 389
Db	456 GLGKLEYLLSHNRLAELPADALGPLQRAFWLDVSHNRLEALPGSLLASLGR 508
RESULT 5	
JC1282	
insulin-like growth factor-binding protein acid labile chain precursor - rat	
C:Species: Rattus norvegicus (Norway rat)	
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004	
C:Accession: JC1282	
R;Dai, J.; Baxter, R.C.	
Biochem. Biophys. Res. Commun. 188, 304-309, 1992	
A:Title: Molecular cloning of the acid-labile subunit of the rat insulin-like growth fac	
A:Reference number: JC1282; MUID:93038676; PMID:1384485	
A:Accession: JC1282	
A:Molecule type: mRNA	
A:Residues: 1-603 <DAI>	
A:Cross-references: UNIPROT:P35859; GB:S46785; NID:g258002; PIDN:AAB23770.2; PID:g570593	
A:Experimental source: liver	
A>Note: the authors translated the codon AAG for residue 63 as Arg, AAA for residue 205	
F:1-27/Domain: signal sequence #status predicted <SIG>	
F:28-603/Product: insulin-like growth factor binding protein, acid labile chain #status	
F:267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>	
Query Match	19.7%; Score 382.5; DB 2; Length 603;
Best Local Similarity	28.7%; Pred. No. 7.5e-19;
Matches	121; Conservative 64; Mismatches 182; Indels 55; Gaps 7;
QY	21 CPKTKCKVVRD----AAQCSCGSAHIAELGLPTNLTHILLFRMDQGILRNHSFGMTV 75
Db	41 CPVACTCSHDDYTDLSVFCSSKNLTHLPD-DIPVSTRALWLDGNNLSSIPSAAFQNLS 99
QY	76 LQRLMLSDSHISAIDPGTFNDLVKLTLRTRNKISRLPRAILDKMVLEQLFLDHNALR 135
Db	100 LDFNLQGSWLRSLEPQALLGLQNLYYLHLERNLRNLAVGLFTHTTFLASLSLSSNLLG 159
QY	136 LLDQNLFQQLRNLOELGLNQQLSFLPANLFSLSRELKLLDLSRNNLTHLPKGLLGAQVK 195
Db	160 RLEGLFQGLSHLWLDNLGWNSLVVLPDPTVFQGLGNLHELVLAGNKLTLYLQPALFCGLGE 219
QY	196 LEKLLYSNQLTSVDSGLLSNLGALTELRLERNHLRSVAPGAF-----DRLG 242
Db	220 LRELDLSRNALRSVKANVFVHLPRQLKLYLDRNLITAVAPGAFGLMKALRWLDLSHNRVA 279
QY	243 NLSS-----LTLSGNLLESPLPALFLHVSSVSRLLTFENPLEELPDVLFGEWAG 291
Db	280 GLMEDTFPGLLGLHVLRLAHNAIASLRPRTFFKDLHFLEELQLGHNRIRQLGERTFEGLGQ 339
QY	292 LRELWNLGTHLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVL----- 344
Db	340 LEVLTLDNQITEVRVGAFSGLFNVAVMNLSGN-CLRSLPERVFQGLDKLHSLHLSCL 398
QY	345 ----ALHT-----NALAELRDDALRGHLRQVSLRHNRLRALPRTLFRNLS 387
Db	399 GHVRLHTFAGLSGLRLFLRDNSISSIEEQSLAGSELLELDTTNRRLTHLPQLFQGLG 458
QY	388 SL 389
Db	459 HL 460
RESULT 6	
JC6128	
insulin-like growth factor binding complex acid labile chain - mouse	
C:Species: Mus musculus (house mouse)	
C:Date: 23-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004	







Db 337 -GAKISLPQAVCDQLPNLQVLDLSYNLLEDL--PSLSGQKQKIDLRHNEIYEIKGST 393

QY 383 FRNLSSL 389

Db 394 FQQLFNL 400

RESULT 9

T13953

MEGF5 protein - rat

N;Alternate names: slit protein homolog

C;Species: Rattus norvegicus (Norway rat)

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C;Accession: T13953

R;Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.

Genomics 51, 27-34, 1998

A;Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs

A;Reference number: Z14126; MUID:98360089; PMID:9693030

A;Accession: T13953

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1523 <NAK>

A;Cross-references: UNIPROT:O88280; EMBL:AB011531; NID:G3449291; PIDN:BAA32461.1; PID:G3449291

C;Genetics:

A;Gene: MEGF5

C;Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein x

Query Match 18.9%; Score 366; DB 2; Length 1523;

Best Local Similarity 20.2%; Pred. No. 3.4e-17;

Matches 120; Conservative 85; Mismatches 180; Indels 210; Gaps 8;

QY 3 RSALLSAVLALLRAQ-PFPCPKTKCVVRDAAQCSGGSVAHIAELGLPTNLTHILLFRMD 61

Db 15 RLALALALASILSGPPAAACPTKCTCSAA-SVDCHGLGLRAVPR-GIPRBAERLDLDRNN 72

QY 62 QGILRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKISRLPRAILDKM 121

Db 73 ITRITKMDFTGLKNLRVLHLEDNQSVIERGAFQDLKQLERLRNKNKLQVLPPELLFQST 132

QY 122 VLLEQLFLDHNALRDLDQNLFOQLRNQLQELGLNQNSFLPANLFSLSRELKLLDLSRNN 181

Db 133 PKLTRLDLSENQIQGIPRKAFRGVTGVKNLQLDNNHISCIEDGAFRALRDLEILTNNNN 192

QY 182 LTHLPKGLLGAQVKLEKLLIYSNQL----- 206

Db 193 ISRILVTSFNHMPKIRTLRHSHNHLVCDCHLAWLSDWLQRRTIGQFTLCMAPVHLRGFS 252

QY 207 -----TSVDSGLLSNLG-----ALTEL 223

Db 253 VADVQKEYVCPGPHSEAPACNANSLCPSACSCSNNIVDCRGKGLTEIPANLPEGIVEI 312

QY 224 RLERNHLRSVAPGAFDRGLGNLSSLTLSGNLLESPPALFLHVSSVSRLTLFENPLEELPD 283

Db 313 RLEQNSIKSI PAGAFIQYKKLRIDISKNQISDIAPDAFQGLKSLTSLVLYGNKITEIPK 372

QY 284 VLFGEAGLRELWNLGTHLSTLPAAAFRNLSGL-----QTL 319

Db 373 GLFDGLVSLQLLLNANKINCLRVNTFQDLQNLNLSLYDNKLQTSKGLFAPLQSIQTL 432

QY 320 GLTRNP----- 325

Db 433 HLAQNPFVCDCHLKLWADLYQDNPIETSGARCSSPRRLANKRISQIKSKKFRCSGSEDYR 492

QY 326 -----RLSALP-----RGV 334

Db 493 NRFSECFMDLVCEKRCRCEGTIVDCSNQKLSRIPSHLPEYTTDLRLNDNDIAVLEATGI 552

QY 335 FQGLRELRLVLAHNTNALAELRDDALRGHLRQVSLRHNRLRALPRTLFRNLSSL 389

Db 553 FKKLPNLRKINLSNNRIKEVREGAFDGAAGVQELMLTGNQLETMTHGRMFRGLSGL 607

RESULT 10

T13852

gene wheeler protein - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C;Accession: T13852

R;Eldon, E.; Kooyer, S.; D'everlyn, D.; Duman, M.; Lawinger, P.; Botas, J.; Bellen, H.

Development 120, 885-899, 1994

A;Title: The Drosophila 18 wheeler is required for morphogenesis and has striking simila

A;Reference number: Z17796; MUID:95324375; PMID:7600965

A;Accession: T13852

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1389 <ELD>

A;Cross-references: UNIPROT:Q24591; EMBL:L23171; NID:G415682; PID:G1019104; PIDN:AAA7920

C;Genetics:

A;Gene: wheeler

A;Cross-references: FlyBase:FBgn0004364

Query Match 17.8%; Score 345; DB 2; Length 1389;

Best Local Similarity 30.4%; Pred. No. 8.6e-16;

Matches 113; Conservative 57; Mismatches 150; Indels 52; Gaps 8;

QY 65 LRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKI----- 110

Db 136 LHGQSFQGLKELSELHLGDNRIQLPEGVWCWCMPSQLNLNTQNRIRSAEFLGFSEKLCA 195

QY 111 -SRLPRA-----IID-----KWLLEQLFLDHNALRDLDQNLF 142

Db 196 GSALSANGAVSGSELTLDVSFNELRSLPDWAGASRLRLQTLSLQHNNTSLAPNAL 255

QY 143 QQLRNQLQELGLNQNSFLPANLFSLSRELKLLDLSRNNLTHLPKGLLGAQVKLEKLL-- 200

Db 256 AGLSSLRVLNISYNHLVSLPSEAFAGNKELRHQLQGNLDLYELPKGLLH---RLEQLLVL 312

QY 201 -LYSNQLTS--VDSGLLSNLGALTELRLERNHLRSVAPGAFDRIGNLSSLTLSGNLLES 257

Db 313 DLSCGNQLTSHVDNSTFAGLRILVLNLSNNALTRIGSKTFKELYFLQILDMRNNNSIGHI 372

QY 258 PPALFLHVSSVSRLTLFENPLEELPDVLFGEAGLRELWNLGTHLSTLPAAAFRNLSGLQ 317

Db 373 BEGAFLPLYNLHTNLAEANRLHTLDNRIFENGLYVLTKLTNNNLVSIVESQAFRNCSDLK 432

QY 318 TLGLTRNPRLSALPRGVFQGLRELRLVLALHTNALAELRDDALRGHLRQVSLRHNRLRA 377

Db 433 ELDLSSN-QLTEVPEAV-QDLSMLKTLDLGENQISEFKNTFRNLNQLTGLRLIDNRIGN 490

QY 378 LPRTLFRNLSSL 389

Db 491 ITVGMFQDLPRL 502

RESULT 11

JE0176

orphan G protein-coupled receptor precursor - human

C;Species: Homo sapiens (man)

C;Date: 03-Jul-1998 #sequence\_revision 10-Jul-1998 #text\_change 09-Jul-2004

C;Accession: JE0176

R;McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.

Biochem. Biophys. Res. Commun. 247, 266-270, 1998

A;Title: Identification and cloning of an orphan G protein-coupled receptor of the glyco

A;Reference number: JE0176; MUID:98308104; PMID:9642114

A;Accession: JE0176

A;Molecule type: mRNA

A;Residues: 1-907 <MCD>

A;Cross-references: UNIPROT:O75473; GB:AF062006; NID:G3366801; PIDN:AAC28019.1; PID:G336

C;Comment: This protein is a receptor for a novel class of glycoprotein ligands.

C;Genetics:

A;Gene: HG38

A;Map position: 12q22-23

F;1-21/Domain: signal sequence #status predicted <SIG>

F;562-583/Domain: transmembrane #status predicted <TM1>

F;594-616/Domain: transmembrane #status predicted <TM2>

F;639-660/Domain: transmembrane #status predicted <TM3>

F;681-701/Domain: transmembrane #status predicted <TM4>  
F;725-744/Domain: transmembrane #status predicted <TM5>  
F;768-791/Domain: transmembrane #status predicted <TM6>  
F;803-824/Domain: transmembrane #status predicted <TM7>

Query Match 17.5%; Score 339; DB 2; Length 907;  
Best Local Similarity 27.7%; Pred. No. 1.3e-15;  
Matches 113; Conservative 69; Mismatches 158; Indels 68; Gaps 10;

QY 13 LLRAQPPCKTKCVVRDAAQCSGSAHI--AELG---LPTNLTHILLFRMDQILRN 67  
Db 30 LLRG---CPTHCHC-----EPDGRMLLRVDCSDGLSELPSNVFTSY----- 70  
QY 68 HSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTTLTRNKISRLPRAILDKMVLLEQL 127  
Db 71 -----LDLSMNNISQLLPNLPSPFLFLEELRLAGNALTYIPKGAFTGLYSKVL 119  
QY 128 FLDHNALRDQNLFFQQLRNLOELGLNQNLQSLPANLFFSSIRELKLLDLSRNNLTHLPK 187  
Db 120 MLQNNQLRHVPTALQNLRLSLQSLRDANHSYVPPSCFSLHSLRHLWLDNALTPIPV 179  
QY 188 GLLGAQVKLEKLIYSNQLTSVDSGLSLNGLALTELRLERNHLRSVAPGAFRLGNLSSL 247  
Db 190 QAFRLSALQMTALNKHIPDYAFGNLSSVLVHLHNNRIHSLGKKCFDGLHSLETL 239  
QY 248 TLSGNLLESPLPAL-----FLHVSSVSRSLTLFENPLEELPDV 284  
Db 240 DLNYYNLDLDEFTAIRTLNLKELGFHNNIRSIPEKAFVGNPSLITIHFDYDNPITQVGRS 299  
QY 285 LFGEMAGRELWNGTHLSTLPAAAFNLSG---LQTLGLTRNPRLSALPRGVFQGLREL 341  
Db 300 AFQHLPELRLTLNGASQIT---EFPDLTGNTANLESILT-GAQISSLRQVCNQLPNL 354  
QY 342 RVLALHTNALAEIRDALRGLHRLQVSLRHNRLRALPRFLRNLSL 389  
Db 355 QVLDLSYNLLEDLPFSFV--CQKLOKIDLRHNEIYEIKVDTFQQLSL 400

RESULT 12.  
Tl3174

gp150 protein - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
C;Accession: Tl3174  
R;Tian, S.S.; Zinn, K.  
J. Biol. Chem. 269, 28478-28486, 1994  
A;Title: An adhesion molecule-like protein that interacts with and is a substrate for a  
A;Reference number: Z17630; MUID:95050638; PMID:7961789  
A;Accession: Tl3174  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1051 <TIA>  
A;Cross-references: UNIPROT:Q24007; EMBL:U15220; NID:G595859; PID:G595860; PIDN:AAA61796  
A;Experimental source: strain Oregon R  
C;Genetics:  
A;Cross-references: FlyBase:FBgn0013272  
A;Map position: 2

Query Match 17.4%; Score 337; DB 2; Length 1051;  
Best Local Similarity 27.1%; Pred. No. 2.1e-15;  
Matches 117; Conservative 65; Mismatches 189; Indels 60; Gaps 10;

QY 16 AQPFPCPKTKCVVRD-----AAQCS-----GGSAHIA----- 44  
Db 228 SEDVPCPRFCQC-ARNVNSYLVAATCSRLDMGIQFGSDITDLVTVNGPKYPILMGPNFF 286  
QY 45 -ELGLPTNLTHILLFRMDQILRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLT 103  
Db 287 QNLGL-KNVASIKIANCTLEYLHAEAFHGLNELYAVNLTVDGLAIINPDTFVGNKKLRML 345  
QY 104 RLTRNKISRLPRA-ILDKMVLLEQLFLDHNAIRDQNLFFQQLRNLOELGLNQNLFLP 162  
Db 346 TISGNDLSVMSSHYLLKSSSIEELDFSRNNLMELNPKAFSHLSNVVYINLSQSLKLP 405

RESULT 13

Tl3887

tlr protein - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: Tl3887  
R;Chiang, C.; Beachy, P.A.  
Mech. Dev. 47, 225-239, 1994  
A;Title: Expression of a novel Toll-like gene spans the parasegment boundary and contrib  
A;Reference number: Z17805; MUID:95151581; PMID:7848870  
A;Accession: Tl3887  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1385 <CHI>  
A;Cross-references: UNIPROT:Q24591; EMBL:S76155; NID:G913247; PID:G913248; PIDN:AAB33383  
C;Genetics:  
A;Cross-references: FlyBase:FBgn0004364  
A;Note: tlr

Query Match 17.4%; Score 337; DB 2; Length 1385;  
Best Local Similarity 29.8%; Pred. No. 3.1e-15;  
Matches 111; Conservative 57; Mismatches 152; Indels 52; Gaps 8;

QY 65 LRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTIRLNKI----- 110  
Db 136 LHGQSFQGLKELSELHGLGDNINIRQLPEGVWCMPSPSLQNLNTQNRIRSAEFLGFSEKLCA 195  
QY 111 -SRLPRA-----ILD-----KMYLLEQLFLDHNAIRDQNL 142  
Db 196 GSALSNANGAVSGGSELOTLDVSPNELRSLPDWAGASRLRLQTLSLQHNNTSLAPNAL 255  
QY 143 QQLRNLOELGLNQNLQSLPANLFFSSIRELKLLDLSRNNLTHLPKGLLGAQVKLEKLL-- 200  
Db 256 AGLSSLRVLNLSYNHLVSLPSEAFAGNKELELHLOQNDLYELPKGLLH---RLEQLLVL 312  
QY 201 -LYSNQLTS--VDSGLSLNGLALTELRLERNHLRSVAPGAFRLGNLSSLTSGNLLES 257  
Db 313 DLSGNQLTDHVDNSTFAGLIRLVLNLSNNAITRIGSKTFKELYFLQILDMRNNNSIGHI 372  
QY 258 PPALFLHVSSVSRSLTLFENPLEELPDVLFEGEMAGRELWNGTHLSTLPAAAFRNLSGLQ 317  
Db 373 EEGAFLPLYNLHTLNLAENRLHTLDNRIFNGLYVLTCLTLNNNLVSIVESQAFRNCSDLK 432  
QY 318 TLGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAEALDDALRGLHRLQVSLRHNRLRA 377  
Db 433 ELDLSSN-QLTEVPEAA-QDLSMLKTLDLGENQISEFKNNTFRNLNQLTGLRLIDNRIGN 490  
QY 378 LPRTLFRNLSSL 389  
Db 491 ITVGMFQDLPL 502

RESULT 14

A58532  
glial cell membrane glycoprotein LIG-1 precursor - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 11-Apr-1997 #sequence\_revision 11-Apr-1997 #text\_change 09-Jul-2004  
C/Accession: A58532  
R/Suzuki, Y.; Sato, N.; Tohyama, M.; Wanaka, A.; Takagi, T.  
J. Biol. Chem. 271, 22522-22527, 1996  
A/Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in  
A/Reference number: A58532; MUID:96394313; PMID:8798419  
A/Accession: A58532  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1091 <SUZ>

A/Cross-references: UNIPROT:P70193; GB:D78572; NID:G1545806; PIDN:BAAL1416.1; PID:G15458  
F/36-61/Domain: proteoglycan amino-terminal homology <PAH>  
F/71-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F/95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F/118-141/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F/142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F/166-189/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F/191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F/214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F/238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F/262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
F/286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>  
F/310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>  
F/334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>  
F/358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>  
F/385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>  
F/409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>  
F/440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>

Query Match 16.9%; Score 327; DB 2; Length 1091;

Best Local Similarity 28.4%; Pred. No. 1.1e-14;

Matches 114; Conservative 78; Mismatches 182; Indels 28; Gaps 11;

QY	6	LLSAVLALL-----RAQ-PFPCPKTKCVVRDAAQCSCGGSVAHIAELGLPT-----N	51
Db	19	LLWLLLLLQWPESAGAQAPRAPCAAACTC-AGNSLDCSGLATLPR-DLPSTWTRSLN	76
QY	52	LTHILLFRMDQGILRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTLRLTRNKIS	111
Db	77	LSYNRLSEIDSA-----AFEDLTNLQEVYLSNELTAI-PSLGTASIGVVSLFLQHNKIL	130
QY	112	RLPRAILDKMVLLEQLFDHNAIRDLDQNLFQQLRNQLQELGLNQNLSTFLPANLFSSL-R	170
Db	131	SVDGSQLKSYLSLEVLDSLNNITEIRSCFPNGRIRELNLASNRISLESFADGLSR	190
QY	171	ELKLLDLSRNNLTHLPKGLIGAQVKLEKLLLYSNQLTSDVSGLLSNLGAITELRLERNHL	230
Db	191	SLLTLRLSKNRITQLPVKAFKLP-RLTQDLNENRIRLIEGLTFQGLDSLEVLRLQRNNI	249
QY	231	RSVAPGAFDRLGNLSSLTLSGNLLESPPALFLHVSSVSRLTLFENPLEELPDVLFGEA	290
Db	250	SRLTDGAFWGLSKMHVLHLEYNLSLVEVNSGSLYGLTALHQHLNNSISRIQRDGSFCQ	309
QY	291	GLRELWNGTHLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHTNA	350
Db	310	KLHELILSFNNLTRLDEESLAELSSLSILRLSHN-AISHIAEGAFKGLKSLRVLDLDHNE	368
QY	351	LAELRDD---ALRGLGHLRQVSLRHNRLRALPRTLFRNLSSL	389
Db	369	ISGTIEDTSGAFTGLDNLKSLTLFGNKIKSVAKRAFSGLESL	410

RESULT 15

B36665  
slit protein 2 precursor - fruit fly (Drosophila melanogaster)  
C/Species: Drosophila melanogaster  
C/Date: 30-Apr-1991 #sequence\_revision 30-Apr-1991 #text\_change 02-Aug-2002  
C/Accession: B36665

R/Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.  
Genes Dev. 4, 2169-2187, 1990

A/Title: slit: an extracellular protein necessary for development of midline glia and c  
A/Reference number: A36665; MUID:91099665; PMID:2176636  
A/Accession: B36665  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-1469 <ROT>  
A/Cross-references: GB:X53959  
C/Genetics:

A/Gene: FlyBase:sl1  
A/Cross-references: FlyBase:FBgn0003425  
C/Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein r  
F/66-91/Domain: proteoglycan amino-terminal homology <PAH1>  
F/101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F/125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F/149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F/173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F/197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F/228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>  
F/288-313/Domain: proteoglycan amino-terminal homology <PAH2>  
F/323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F/347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F/371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F/395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
F/419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>  
F/450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>  
F/512-537/Domain: proteoglycan amino-terminal homology <PAH3>  
F/547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>  
F/572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>  
F/596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>  
F/620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>  
F/651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>  
F/708-733/Domain: proteoglycan amino-terminal homology <PAH4>  
F/743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>  
F/767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>  
F/846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>  
F/1028-1061/Domain: EGF homology <EGF>  
F/1068-1099/Domain: EGF homology <EGF2>  
F/1115-1148/Domain: EGF homology <EGF1>

Query Match 16.5%; Score 319.5; DB 2; Length 1469;

Best Local Similarity 29.3%; Pred. No. 5.4e-14;

Matches 118; Conservative 64; Mismatches 156; Indels 65; Gaps 14;

QY	21	CPKCKCVVRDAAQCSCGGSV-----AHIAELGLP-TNLTHILLFRMDQGILRNHSFS	71
Db	73	CPRVCSCTGLN-VDCSHRGITSVPRKISADVERLELQGNLTF-----VIYETDFQ	121
QY	72	GMTVLQRLMLSDSHISAIDPGTFNDLVKLTLRLTRNKISRLPRAILDKMVLLEQLFLDH	131
Db	122	RLTKLRMLQLTNDQIHTIENSFQDLVSLERLIDISNNVITTVGRRVFKGAQSLRSQLDN	181
QY	132	NALRDLQNLFQQLRNQLQELGLNQNLSTFLPANLFSSLRELKLLDLNRN-----NLTHL	185
Db	182	NQITCLDEHAFKGLVELEILTNNNNLTSLPHNIFGGLRRLRALRLSDNPFACDCHLSWL	241
QY	186	PKGLLGAQVKLEKLLLYS-----NQLTSVD-----SGLLSNLGALTTELRLERN-	228
Db	242	SRFLRSA---TRLAPYTRCQSPSQKQGVADLHDQEFKCSGLTEH--APMECCGAENSC	295
QY	229	-HLRSVAPGAFD-RLGNLSSLTLSGNLLESPPALFLHVSSVSRLTLFENPLEELPDVLF	286
Db	296	FHPCRCADGIVDCREKSLTSVPV-----TLP-----DDTTDVRLEQNFTTELPPKSF	342
QY	287	GEMAGLRELWNGTHLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVAL	346
Db	343	SSFRRLRRIDLNNNISRIADALSGLKQLTTLVLYGN-KIKDLPSPGVFKGLSLRLILL	401
QY	347	HTNALAELRDDALRGLGHLRQVSLRHNRLRALPRTLFRNLSSL	389
Db	402	NANEISCIRKDAFRDLHSLSLSLSLYDNNIQSLANGTFDAMKSM	444

Wed Dec 1 15:24:46 2004

us-09-775-803-12\_copy\_1\_389.rpr

Page 8

Search completed: November 29, 2004, 13:11:29  
Job time : 27 secs